Applying the Explicit Aggregation Algorithm to Heterogeneous Macro Models*

Takeki Sunakawa†

July 2019

Abstract

This paper applies the explicit aggregation (Xpa) algorithm developed by den Haan and Rendahl [2010. Solving the incomplete markets model with aggregate uncertainty using explicit aggregation. Journal of Economic Dynamics and Control, 69–78] to the heterogeneous–firm models of Khan and Thomas (2003, 2008) and the heterogeneous–household models of Krueger et al. (2016). We find the Xpa algorithm is an order of magnitude faster for solving these models than the standard Krusell–Smith (KS) algorithm because it does not need to simulate the distribution of individual capital and productivity when updating the aggregate forecasting rules. However, the simulation results in the Xpa and KS algorithms in terms of both the micro– and macro–level moments are almost identical, even in levels, because of bias correction. The Xpa algorithm also exhibits accuracy comparable with the KS algorithm.

*This paper is originally a part of my dissertation at the Ohio State University. The author is especially grateful to Aubhik Khan and Julia Thomas for their valuable guidance and support, and to Pontus Rendahl for sharing his code for bias correction. The author also appreciates discussions with anonymous referees, Jie April Cai, Grey Gordon, Youngsoo Jang, Heejeong Kim, Timothy Kam, Toshihiko Mukoyama, Taisuke Nakata, Nobuhide Okahata, Tatsuro Senga, Shuhei Takahashi, Stephen Terry, Tomoaki Yamada, and Minchul Yum, as well as comments by seminar participants at Kyoto University, the Ohio State University, the Australian National University, and the Shanghai University of Finance and Economics, and attendees at the 2013 Conference on Computing in Economics and Finance, the 10th Dynare Conference, the 2019 Spring Midwest Macroeconomics Meetings, and the 2019 Conference on Computing in Economics and Finance. The usual disclaimer applies. This research is financially supported by a Japan Society for the Promotion of Science Grant–in–Aid for Scientific Research (KAKENHI) for Young Scientists (B) Project No. 15K20918 and the Seimeikai Foundation. The programming code used in the paper is publically available at `https://github.com/tkksnk/Xpa`.

†Kobe University Center for Social Systems Innovation. E–mail: takeki.sunakawa@gmail.com.

Keywords: Projection methods, Heterogeneous agents, Explicit aggregation.

JEL codes: C63; D52

1 Introduction

Solving heterogeneous–agent macro models accurately and efficiently is more than ever an important research agenda. Some recent studies revisit the importance of understanding the effect of micro–level heterogeneity on aggregate macro dynamics (Ahn et al., 2018; Krueger et al., 2016). For this reason, as more micro–level panel data consistent with macro aggregates become available, structural estimations of heterogeneous–agent models using both micro– and macro–level data are of increasing interest. Accordingly, a faster and accurate computation method for heterogeneous–agent models is necessary. den Haan and Rendahl (2010) develop the explicit aggregation (Xpa) algorithm to solve the heterogeneous–household model studied in Krusell and Smith (1998). A special issue in the *Journal of Economic Dynamics and Control* (den Haan et al., 2010) considers different algorithms and shows that the Xpa algorithm can solve the model more efficiently than the standard Krusell–Smith (KS) algorithm (Maliar et al., 2010).

In this paper, the Xpa algorithm originally proposed by den Haan and Rendahl (2010) is applied to the heterogeneous–firm models studied in Khan and Thomas (2003, 2008) (hereafter KT) and to the heterogeneous–household models examined in Krueger et al. (2016) (hereafter KMP). The latter application is especially interesting as they demonstrate that micro–level heterogeneity matters for the aggregate macro dynamics when the model is matched with the micro–level wealth distribution, which is not the case in Krusell and Smith (1998). Indeed, the Xpa algorithm is an order of magnitude faster than the KS algorithm for solving these models because it does not need to simulate the distribution of individual capital and productivity when updating the aggregate forecasting rules. The Xpa algorithm is also easy to implement as we merely evaluate the individual decision rules *as if* the amount of individual capital is equal to that of the aggregate capital, e.g., $k = K_t$, especially in the case without idiosyncratic shocks.

We first show how to apply the Xpa algorithm to the KT (2003) model without idiosyncratic productivity shocks (see also Thomas, 2002). Given Jensen's inequality, bias correction using steady–state information is the key to accuracy in terms of forecasting errors such as the den Haan (2010) statistics. We find the simulation results using the Xpa and KS algorithms are almost

identical in terms of both the micro– and macro–level moments, even in levels, because of the bias correction. Further, the Xpa algorithm is nearly equally as accurate as the KS algorithm. We also present how the Xpa algorithm can apply to the KT (2008) model with idiosyncratic productivity shocks, where the Xpa algorithm has den Haan errors at most only twice as large as those using the Krusell–Smith algorithm. We accomplish this developing a novel approach we call ϵ–indexed aggregation to deal with the idiosyncratic shocks.[1]

In terms of related work, Terry (2017) applies different algorithms to solve the KT (2008) model and asserts that the Xpa algorithm is indeed faster, but less accurate, than other algorithms employing projection methods. However, that implementation of the Xpa algorithm applies a naive aggregation.[2] Instead, in the present paper, we argue that the ϵ–indexed aggregation provides much more accurate solutions by a factor of three or four more than what Terry (2017) reports for the den Haan errors.

We then show that the Xpa algorithm also works in a KMP (2016) environment in which heterogeneity matters. The ϵ–indexed aggregation is modified so that we can deal with exogenously time–varying employment (Krusell and Smith, 1998). This is a non–trivial extension as the dimension of idiosyncratic shocks is made much larger to include not only employment status, but also earnings risk and discount factor shocks, in KMP's benchmark model.[3] Once again, the simulation results using the Xpa and KS algorithms are very close and the accuracy in the forecasting errors only slightly worsens with the Xpa algorithm. In our computation, we solve the benchmark model in KMP in seconds with the Xpa algorithm using Fortran and OpenMP.

Krusell and Smith (1998) first develop a simulation–based method for solving an incomplete–market heterogeneous–household model with aggregate uncertainty. A computational difficulty

[1]The ϵ–indexed aggregation evaluates the individual decision rules at $k = K_t(\epsilon_i)$, where $K_t(\epsilon_i)$ is the total amount of capital held by individuals that have a common idiosyncratic shocks ϵ_i. Given that the measure of households $\mu_t(\epsilon_i, k)$ can be decomposed into $\mu_{kt}(k|\epsilon_i)\phi(\epsilon_i)$, we can use $K_t(\epsilon_i) = \int k\mu_{kt}(k|\epsilon_i)dk$ instead of $K_t = \sum_i \int k\mu_t(\epsilon_i, k)$ to evaluate the individual decision rules and take a weighted average of the evaluated functions with $\phi(\epsilon_i)$ being the weight. As we know only the values of K_t as grid points, we approximate $K_t(\epsilon_i) \approx \psi(\epsilon_i)K_t$ where $\psi(\epsilon_i) = K_{ss}(\epsilon_i)/K_{ss}$ by using steady–state information.

[2]The Xpa algorithm is first adapted and applied to the heterogeneous–firm models by Sunakawa (2012), which also uses naive aggregation.

[3]den Haan and Rendahl (2010) include each ϵ–indexed capital in the indivisual decision rule as a state variable. This approach is very costly or infeasible as the number of grid points for ϵ_i increases because we need more dimensions of the state space.

arises from the fact that the state of the economy is the whole distribution of individual capital and productivity. To address this, they assume that *approximate aggregation* (Krusell and Smith, 1998; Young, 2005) holds, such that the entire state of the economy is approximated by the mean of the distribution. They then update the forecasting rules for individual households to infer the state of the economy by simulating the model for a long time, which is time–consuming.[4]

In contrast, Reiter (2009) proposes a way to handle the whole distribution without assuming the aggregate approximation.[5] The distribution is approximated by the non–stochastic method with histogram (Young, 2010). Then, the equilibrium dynamics with aggregate shocks are approximated at the first order by linearizing the model around the deterministic steady state. The method can then handle hundreds of state variables because of such linearization.[6] Although Reiter's (2009) method is also much faster than the standard KS algorithm, as is the Xpa algorithm, the method is based on first–order perturbation and cannot consider the higher–order effects of aggregate uncertainty as certainty equivalence holds.[7]

The reminder of the paper proceeds as follows. Section 2 presents the heterogeneous–firm models in KT (2003; 2008). Section 3 explains the KS and Xpa algorithms for each of the cases without and with idiosyncratic shocks. Especially in the latter, we introduce ϵ–indexed aggregation to deal with the idiosyncratic shocks. We also show how to adjust the bias in levels from Jensen's inequality by following den Haan and Rendahl (2010). Section 4 analyzes the numerical results of the heterogeneous–firm models with the calibrated parameters in KT (2003; 2008) and Section 5 presents the heterogeneous–household models in KMP (2016) and provides the calibration and numerical results. We find that in both cases of the heterogeneous–firm and heterogeneous–household models, the Xpa algorithm is much faster and comparably accurate to the KS algorithm. Section

[4] Alternatively, Algan et al. (2008, 2010) and Winberry (2018) show how to solve the model by parametrizing the distribution. Gordon (2011) proposes a way to approximate the whole distribution with hundreds of state variables (i.e., bins in a histogram) with sparse grid.

[5] Reiter's (2010b) backward induction method can also be applied to models in which approximate aggregation does not hold. The method is applied to stochastic overlapping generations models with aggregate uncertainty by Khan (2017); Kim (2018). Okahata (2018) demonstrates the method can also merge with the continuous–time methods in Ahn et al. (2018).

[6] Reiter (2010a); Ahn et al. (2018) further develop a method to reduce the dimension of the state space by projecting the distribution onto principal components. See also Bayer and Luetticke (2018). Childers (2018) suggests a distinct approach using linearization.

[7] Terry (2017) compares Reiter's method and the standard KS method in solving variants of the KT (2008) model.

6 concludes. Additional details of the computation are in the Appendix.

2 Khan–Thomas models

In this section, we briefly explain the heterogeneous–firm models in KT (2003; 2008) and refer the interested readers to the original articles for more detailed discussion. In brief, KT (2008) provides two extensions to KT (2003), namely, idiosyncratic productivity shocks and constrained investment without fixed costs, so as to match the micro–level moments of the investment–capital ratio in the model and the data.

In this economy, there is a continuum of firms. Each firm produces output using the production function: $y = \epsilon_i z_t F(k, n)$, where k is the capital stock, n is the labor input, $\epsilon_i \in \{\epsilon_1, ..., \epsilon_{N_\epsilon}\}$ is the idiosyncratic firm–specific productivity and $z_t \in \{z_1, ..., z_{N_z}\}$ is the aggregate productivity common across firms. Each of ϵ_i and z_t follows a Markov chain. $F(k, n)$ is a common production function that exhibits decreasing returns to scale. A random fixed cost ξ is incurred on adjusting individual capital in each period.

A firm is defined by the individual state variables (ϵ_i, k, ξ). The aggregate state of the economy is given by (z_t, μ_t), where μ_t is the distribution of individual capital and productivity (ϵ_i, k) held by each firm, which is an infinitesimal object.[8] Taking as given the shadow price of the household's utility p_t and the wage rate w_t, the firm chooses labor input n and the next period's capital k'.

The firm's choice on labor is static. The firm chooses n so as to maximize the current profit

$$\pi(\epsilon_i, k; z_t, \mu_t) = \max_n \{\epsilon_i z_t F(k, n) - w_t n\} + (1 - \delta)k, \tag{1}$$

where $w_t = \Gamma_w(z_t, \mu_t)$ is the wage rate, which is a function of the aggregate state variables. Note that the profit includes undepreciated capital $(1 - \delta)k$ with the depreciation rate of $\delta \in (0, 1]$ from the previous period.

The firm's choice on capital is dynamic and discrete. That is, if the firm pays a fixed cost (proportional to the wage rate) ξw_t, the firm can choose any level of capital in the next period, k'.

[8] We use μ_t and $\mu_t(\epsilon_i, k)$ interchangeably for the distribution, which is also the measure of firms with (ϵ_i, k).

If not, the firm can choose k' within a limited range of $\Omega(k)$.[9] The firm's value by choosing the next period's capital k' is

$$\begin{aligned}
& \tilde{v}_0(\epsilon_i, k, \xi; z_t, \mu_t) \\
= \; & \pi(\epsilon_i, k; z_t, \mu_t) + \max\{-\xi w_t \\
& + \max_{k'>0}\left\{-\gamma k' + \beta \sum_{\epsilon'|\epsilon_i} \pi_\epsilon(\epsilon'|\epsilon_i) \sum_{z_{t+1}|z_t} \pi_z(z_{t+1}|z_t) \frac{p_{t+1}}{p_t} \tilde{v}(\epsilon', k'; z_{t+1}, \mu_{t+1})\right\}, \\
& \max_{k'\in\Omega(k)}\left\{-\gamma k' + \beta \sum_{\epsilon'|\epsilon_i} \pi_\epsilon(\epsilon'|\epsilon_i) \sum_{z_{t+1}|z_t} \pi_z(z_{t+1}|z_t) \frac{p_{t+1}}{p_t} \tilde{v}(\epsilon', k'; z_{t+1}, \mu_{t+1})\right\}\Bigg\}, \qquad (2)
\end{aligned}$$

where $p_t = \Gamma_p(z_t, \mu_t)$ is the (shadow) price and $\mu_{t+1} = \Gamma_\mu(z_t, \mu_t)$ is the forecasted distribution in the next period. $\pi_\epsilon(\epsilon'|\epsilon_i)$ ($\pi_z(z_{t+1}|z_t)$) is the conditional probability of the next period's idiosyncratic (aggregate) productivity given the current period's productivity. $\tilde{v}(\epsilon_i, k; z_t, \mu_t) = \int_0^{\bar{\xi}} \tilde{v}_0(\epsilon_i, k, \xi; z_t, \mu_t) d\xi$ is the ex ante value of the firms that commonly have (ϵ_i, k) before they draw a random fixed cost ξ. Each firm discounts the future by a common stochastic discount factor $\beta p_{t+1}/p_t$, whereas $\beta \in (0, 1)$ is the deterministic discount factor. γ is the deterministic trend of the economy.

We define $v(\epsilon_i, k; z_t, \mu_t) \equiv p_t \tilde{v}(\epsilon_i, k; z_t, \mu_t)$. By multiplying p_t on both sides of (2) and integrating the equation over ξ, we have a transformed Bellman equation that is recursive in $v(\epsilon_i, k; z_t, \mu_t)$:

$$\begin{aligned}
v(\epsilon_i, k; z_t, \mu_t) &= p_t \pi(\epsilon_i, k; z_t, \mu_t) \\
&- p_t \int_0^{\hat{\xi}(\epsilon_i, k; z_t, \mu_t)} \xi G'(\xi) d\xi w_t \\
&+ \alpha(\epsilon_i, k; z_t, \mu_t) E_0(\epsilon_i; z_t, \mu_t) + (1 - \alpha(\epsilon_i, k; z_t, \mu_t)) E_1(\epsilon_i, k; z_t, \mu_t), \qquad (3)
\end{aligned}$$

[9] If $\Omega(k) = (1-\delta)k$ is a point set, firms hold the current level of capital after depreciation.

where

$$E_0(\epsilon_i; z_t, \mu_t) = \max_{k'>0} \left\{ -\gamma p_t k' + \beta \sum_{\epsilon'|\epsilon_i} \pi_\epsilon(\epsilon'|\epsilon_i) \sum_{z_{t+1}|z_t} \pi_z(z_{t+1}|z_t) v(\epsilon', k'; z_{t+1}, \mu_{t+1}) \right\}, \quad (4)$$

$$E_1(\epsilon_i, k; z_t, \mu_t) = \max_{k' \in \Omega(k)} \left\{ -\gamma p_t k' + \beta \sum_{\epsilon'|\epsilon_i} \pi_\epsilon(\epsilon'|\epsilon_i) \sum_{z_{t+1}|z_t} \pi_z(z_{t+1}|z_t) v(\epsilon', k'; z_{t+1}, \mu_{t+1}) \right\}. \quad (5)$$

If the firm chooses to pay the fixed cost, it selects $k^*(\epsilon_i; z_t, \mu_t)$ regardless of current capital k and the value of adjustment is $E_0(\epsilon_i; z_t, \mu_t)$. If the plant chooses not to pay the fixed cost, the next period's capital is $k^c(\epsilon_i, k; z_t, \mu_t)$ and the value of non–adjustment is $E_1(\epsilon_i, k; z_t, \mu_t)$. $\hat{\xi}(\epsilon_i, k; z_t, \mu_t) = \left(E_0(\epsilon_i; z_t, \mu_t) - E_1(\epsilon_i, k; z_t, \mu_t)\right) / \left(w_t p_t\right)$ is the threshold level of the fixed cost at or below which the firm chooses to pay the fixed cost as $E_0(\epsilon_i; z_t, \mu_t) \geq E_1(\epsilon_i, k; z_t, \mu_t)$ holds. G is the cumulative density function of the distribution $G : [0, \bar{\xi}] \to [0, 1]$ and $\alpha(\epsilon_i, k; z_t, \mu_t) = G\left(\hat{\xi}(\epsilon_i, k; z_t, \mu_t)\right)$ is the probability of adjusting capital.

Given the decision rules for adjusters and non–adjusters, the *ex post* firm–level decision rule for capital is

$$k' = K(\epsilon_i, k, \xi; z_t, \mu_t) = \begin{cases} k^*(\epsilon_i; z_t, \mu_t), & \text{if } \xi \leq \min\{\bar{\xi}, \hat{\xi}(\epsilon_i, k; z_t, \mu_t)\}, \\ k^c(\epsilon_i, k; z_t, \mu_t), & \text{if } \xi > \min\{\bar{\xi}, \hat{\xi}(\epsilon_i, k; z_t, \mu_t)\}. \end{cases}$$

By convexifying $K(\epsilon_i, k, \xi; z_t, \mu_t)$ over ξ among the firms that commonly have k, we obtain the *ex ante* firm–level decision rule for capital:

$$g_k(\epsilon_i, k; z_t, \mu_t) = \alpha(\epsilon_i, k; z_t, \mu_t) k^*(\epsilon_i; z_t, \mu_t) + (1 - \alpha(\epsilon_i, k; z_t, \mu_t)) k^c(\epsilon_i, k; z_t, \mu_t). \quad (6)$$

In addition, the firm–level decision rule for output is given by

$$g_y(\epsilon_i, k; z, \mu) = z_t F(\epsilon_i, k, n^*(\epsilon_i, k; z_t, \mu_t)), \quad (7)$$

where $n^*(\epsilon_i, k; z_t, \mu_t)$ is the optimal level of labor input that the firms choose by solving the intratemporal problem (1).

The individual decision rules are aggregated as

$$Y_t = \sum_i \int g_y(\epsilon_i, k; z_t, \mu_t)\mu_t(\epsilon_i, k)dk, \tag{8}$$

$$K_{t+1} = \sum_i \int g_k(\epsilon_i, k; z_t, \mu_t)\mu_t(\epsilon_i, k)dk, \tag{9}$$

where $\mu_t(\epsilon_i, k)$ is the measure of firms with (ϵ_i, k). We normalize the distribution so that $\sum_i \int \mu_t(\epsilon_i, k)dk = 1$ holds. Having Y_t and K_{t+1} at hand, the resource constraint implies

$$C_t = Y_t + (1-\delta)K_t - K_{t+1}$$

where C_t is the aggregate consumption and $K_t = \sum_i \int k\mu_t(\epsilon_i, k)dk$ is the aggregate capital. The labor market clears such that

$$N_t = \sum_i \int \left[n^*(\epsilon_i, k, z_t, \mu_t) + \int_0^{\hat{\xi}(\epsilon_i, k; z, \mu)} \xi G'(\xi)d\xi \right] \mu_t(\epsilon_i, k)dk$$

holds, where N_t is the aggregate hours worked.

Finally, the prices and wages are functions of the aggregate state of the economy because the household's consumption–saving and working decisions have effects on the prices and wages in a general equilibrium. Let $U(C_t, 1-N_t) = \log C_t + \eta(1-N_t)$ be the household's utility function period by period. Then the prices and wages are given by $p_t = D_1U(C_t, 1-N_t) = 1/C_t$ and $w_t = D_2U(C_t, 1-N_t)/D_1U(C_t, 1-N_t) = \eta/p_t$ as a result of the household's optimization.

A recursive competitive equilibrium is defined so as to satisfy (i) the household's optimality, (ii) each firm's optimality, (iii) feasibility (i.e., markets clearing), and (iv) consistency between the individual decision rules and the aggregate forecasting rules.

Regarding the feasibility stated in the definition above, the individual decision rules $g_y(\epsilon_i, k; z_t, \mu_t)$ and $g_k(\epsilon_i, k; z_t, \mu_t)$ are implicit functions of the prices p_t and wages $w_t = \eta/p_t$. For a conjectured price $\tilde{p}_t$, (4) and (5) are solved for the individual decision rules. Then (8) and (9) yield the aggregate variables Y_t, K_{t+1}, and the updated price $p_t = C_t^{-1}$. There is a mapping denoted by $p_t = F(\tilde{p}_t)$

from the conjectured price to the updated price. To obtain the market–clearing price, a fixed point problem $p_t^* = F(p_t^*)$ is solved.[10]

Regarding consistency, each of the Xpa or KS algorithms achieves this in a different manner, as shown in the following section.

3 Two algorithms

In this section, we explain the two algorithms used in this study, namely, the Xpa algorithm developed by den Haan and Rendahl (2010) and the standard KS algorithm. The individual decision rules are consistent with the forecasting rules for the aggregate state of the economy. The Xpa and KS algorithms achieve consistency in a different manner. The only difference in these algorithms is their way of aggregating the individual decision rules to obtain the forecasting rules. For the most part, the Xpa algorithm is easier to implement as it does not require simulations, but it is not straight–forward to apply the method to models with idiosyncratic shocks. For this reason, we present a novel approach we call ϵ–indexed aggregation to deal with idiosyncratic shocks.

3.1 Without idiosyncratic shocks

We first consider the case without idiosyncratic shocks as in KT (2003). We remove idiosyncratic shocks ϵ_i and constrain investment by setting $\Omega(k) = (1-\delta)k$ (i.e., no adjustment is allowed without paying the fixed cost) in the model presented in the previous section. We also assume that the approximate aggregation of Krusell and Smith (1998) and Young (2005) holds. That is, the mean of the distribution at the beginning of period t, $K_t = \int k\mu_t(k)dk$, is sufficient for individuals to forecast the state of the economy. Therefore, K_t enters into the individual decision rules instead of μ_t. Taking the individual decision rules with the aggregate approximation $y = g_y(k; z_t, K_t)$ and $k' = g_k(k; z_t, K_t)$ as given, we obtain the aggregate output and the next period's aggregate capital

[10] C_t and N_t are unpredetermined variables. Therefore, the market–clearing prices are solved to ensure the feasibility. This is also the case in Chang and Kim (2007, 2014); Takahashi (2014); Jang et al. (2018). In contrast, the prices and wages are functions of only predetermined variables in Krusell and Smith (1998).

by

$$Y_t = \int g_y(k; z_t, K_t)\mu_t(k)dk, \tag{10}$$

$$K_{t+1} = \int g_k(k; z_t, K_t)\mu_t(k)dk. \tag{11}$$

We also obtain $C_t = Y_t + (1-\delta)K_t - K_{t+1}$ and $p_t = C_t^{-1}$. As discussed earlier, we impose market clearing.

In the KS algorithm, we simulate the model for a long time to obtain the sequence of $\{z_t, K_t, p_t\}_{t=1}^T$ taking the sequence of exogenous aggregate productivity shock $\{z_t\}$ and the initial value of K_1 as given. Using the sequence obtained, we construct a subset of the sequence sorted by the realization of z_t for each $z_i \in \{z_1, ..., z_{n_z}\}$ and estimate the forecasting rule by ordinary least squares in the form of

$$\log K_{t+1} = b_{K,0}(z_i) + b_{K,1}(z_i)\log K_t,$$
$$\log p_t = b_{p,0}(z_i) + b_{p,1}(z_i)\log K_t,$$

for each $i = 1, ..., n_z$. We denote them by $K_{t+1} = \Gamma_K^{\mathrm{KS}}(K_t, z_t)$ and $p_t = \Gamma_p^{\mathrm{KS}}(K_t, z_t)$.

In the Xpa algorithm, instead of simulating the model, we just move the integral over k into the individual decision rule

$$\begin{aligned} K_{t+1} &= \int g_k(k; z_t, K_t)\mu_t(k)dk, \\ &\approx g_k(\int k\mu_t(k)dk; z_t, K_t), \\ &= g_k(K_t; z_t, K_t). \end{aligned}$$

Note that when we solve for the individual decision rule, K_t is given as a grid point instead of the integral of the distribution. We immediately obtain the forecasting rule for aggregate capital in the Xpa algorithm $K_{t+1} = \Gamma^{\mathrm{Xpa}}(K_t, z_t) \equiv g(K_t; z_t, K_t)$ by evaluating the individual decision rule $g_k(k; z_t, K_t)$ at $k = K_t$. We also obtain $Y_t = g_y(K_t; z_t, K_t)$, which implies the forecasting rule for

the shadow price of the household's utility by $C_t = Y_t + (1-\delta)K_t - K_{t+1}$ and $p_t = C_t^{-1}$. We again consider market clearing.

Bias correction Because of Jensen's inequality, the Xpa algorithm yields biases in Y_t and K_{t+1}. The individual decision rules $g_y(k; z_t, K_t)$ and $g_k(k; z_t, K_t)$ must be linear at $k = K_t$ for the Xpa algorithm to have the same aggregate variables as in (10) and (11) obtained in the KS algorithm.

den Haan and Rendahl (2010) suggested a way to correct these biases in Y_t and K_{t+1}. Let $K_{ss} = \int \bar{g}_k(k)\bar{\mu}(k)dk$ be the level of aggregate capital, which is obtained by the steady–state individual decision rule $\bar{g}_k(k)$ and the stationary distribution $\bar{\mu}(k)$. Also let $\tilde{K}_{ss} = \bar{g}_k(\int k\bar{\mu}(k)dk) = \bar{g}_k(K_{ss})$ be the decision rule $\bar{g}_k(k)$ evaluated at $k = K_{ss}$. Then a time–invariant number for bias correction is given by $\zeta_k \equiv K_{ss} - \tilde{K}_{ss} = K_{ss} - \bar{g}_k(K_{ss})$, which measures Jensen's inequality in the steady state. We do the same for output and obtain $\zeta_y \equiv Y_{ss} - \tilde{Y}_{ss} = Y_{ss} - \bar{g}_y(K_{ss})$. We then add the bias–correction terms to the forecasting rules in the Xpa algorithm, that is,

$$K_{t+1} = g_k(K_t; z_t, K_t) + \zeta_k,$$
$$Y_t = g_y(K_t; z_t, K_t) + \zeta_y.$$

Figure 1 illustrates the shape of the individual decision rule $g_k(k; z_t, K_t)$ evaluated at $k = K_t$. The figure is a typical example of (S, s)–type capital adjustment. That is, only when the current level of individual capital is sufficiently far from the target level will the firm choose to adjust capital; otherwise, the firm will continue holding the current level of capital (after capital depreciation). Even after convexifying the ex post firm–level decision rule over random fixed costs, the ex ante firm–level decision rule exhibits considerable nonlinearity. This is quite different from the individual decision rules in the household consumption–saving problem in Krusell and Smith (1998), which are almost linear save points near the origin where poor households face borrowing constraints.

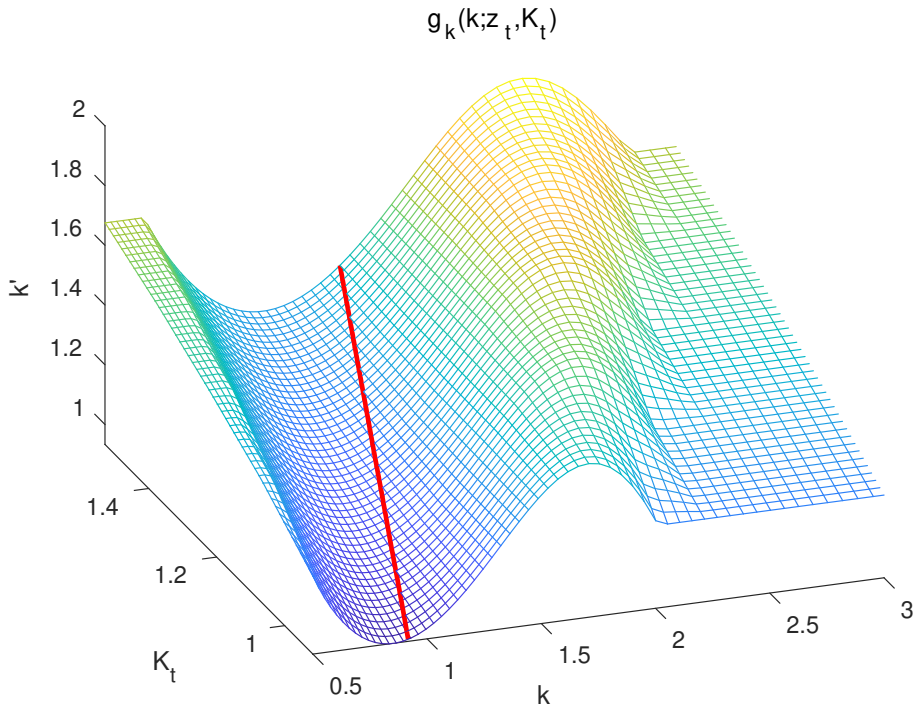
Even though the individual decision rule $g_k(k; z_t, K_t)$ is nonlinear in k, when it is evaluated at $k = K_t$, the resulting function $g_k(K_t; z_t, K_t)$ is linear in K_t.[11] The red line in the figure depicts

[11] This result may suggest that the Xpa algorithm works when the forecasting rule obtained by the KS algorithm exhibits small nonlinearity, e.g., a high R–squared.

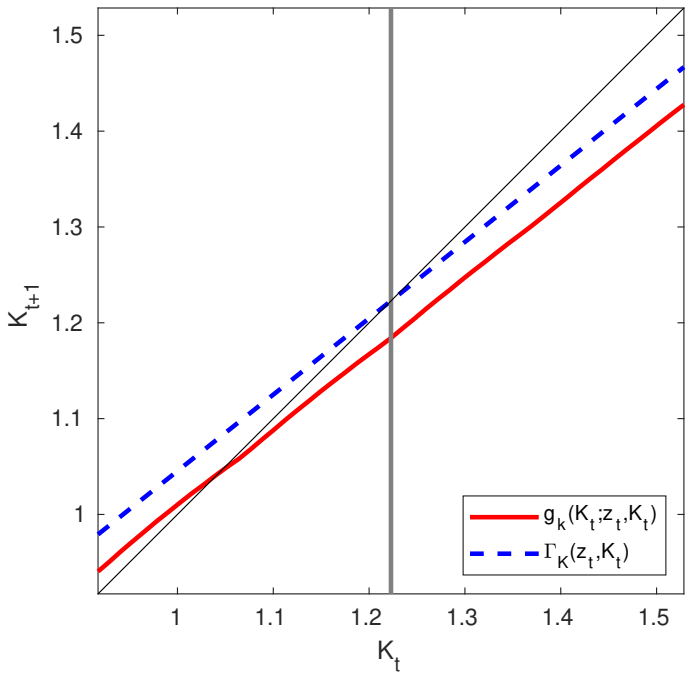
this relationship. Given Jensen's inequality, there is a downward bias in $g_k(K_t; z_t, K_t)$ compared with $\Gamma_K(z_t, K_t)$. Note that $g_k(k; z_t, K_t)$ is convex in k taking (z_t, K_t) as fixed. We can correct for this bias using the steady–state counterparts, $\bar{g}_k(k)$. After bias correction, we can see that the intersection of $\Gamma_K(z_t, K_t)$ and the 45–degree line (in black) corresponds to the deterministic steady state of K_{ss} (the vertical gray line).[12]

[12] Terry (2017) uses a different method for correcting biases. The intercepts in the forecasting rules, $\log K_{t+1} = b_{K,0}(z_i) + b_{K,1}(z_i) \log K_t$, are adjusted so that the steady state implied by the forecasting rules $\exp\left(\frac{b_{K,0}(z_i)}{1-b_{K,1}(z_i)}\right)$ coincides with the deterministic steady state K_{ss} for each level of z_i. In the proposed method, we undertake such a correction using the individual decision rules in the steady state.

Figure 1: Individual decision rules and forecasting rules.



Notes: The x–axis is the current period's individual capital k, the y–axis is the aggregate capital K_t, and the z–axis is the next period's individual capital k'. $z_t = 1.0$.



Notes: The x–axis is the current period's aggregate capital K_t and the y–axis is the next period's aggregate capital K_{t+1}.

3.2 With idiosyncratic shocks

When we allow for idiosyncratic shocks, aggregate capital in the next period (we omit aggregate output, but the same discussion applies) is given by

$$K_{t+1} = \sum_i \int g_k(\epsilon_i, k; z_t, K_t)\mu_t(\epsilon_i, k)dk,$$

The measure of (ϵ_i, k) can be decomposed into

$$\mu_t(\epsilon_i, k) = \mu_{kt}(k|\epsilon_i)\phi(\epsilon_i),$$

where $\phi(\epsilon_i) = \int \mu_t(\epsilon_i, k)dk$ is the marginal distribution of ϵ_i assumed to be a time–invariant measure of ϵ_i and $\mu_{kt}(k|\epsilon_i) = \frac{\mu_t(\epsilon_i,k)}{\int \mu_t(\epsilon_i,k)dk}$ is the conditional distribution of $\mu_t(\epsilon_i, k)$.[13] Then, the aggregation in the Xpa algorithm becomes

$$\begin{aligned} K_{t+1} &= \sum_i \int g_k(\epsilon_i, k; z_t, K_t)\mu_{kt}(k|\epsilon_i)\phi(\epsilon_i)dk, \\ &\approx \sum_i g_k(\epsilon_i, \int k\mu_{kt}(k|\epsilon_i)dk; z_t, K_t)\phi(\epsilon_i), \\ &= \sum_i g_k(\epsilon_i, K_t(\epsilon_i); z_t, K_t)\phi(\epsilon_i), \end{aligned}$$

where $K_t(\epsilon_i) \equiv \int k\mu_{kt}(k|\epsilon_i)dk$ is the amount of capital indexed by ϵ_i at the beginning of period t. Note that $K_t = \sum_i \int k\mu_t(k, \epsilon_i)dk = \sum_i K_t(\epsilon_i)\phi(\epsilon_i)$ holds.

ϵ–indexed aggregation We wish to use $K_t(\epsilon_i)$ to evaluate the individual decision rule. However, we know the value of K_t only as a grid point. Therefore, we apply so–called ϵ–indexed aggregation as follows. $K_t(\epsilon_i)$ is not on grid points and we evaluate the individual decision rule at $K_t(\epsilon_i) = \psi(\epsilon_i)K_t$ where

$$\psi(\epsilon_i) = \frac{K_{ss}(\epsilon_i)}{K_{ss}} = \frac{K_{ss}(\epsilon_i)}{\sum_i K_{ss}(\epsilon_i)\phi(\epsilon_i)},$$

[13] Note that its integral over k is normalized to one, i.e., $\int \mu_{kt}(k|\epsilon_i)dk = \frac{\int \mu_t(\epsilon_i,k)dk}{\int \mu_t(\epsilon_i,k)dk} = 1$.

$K_{ss}(\epsilon_i) = \int k\bar{\mu}_k(k|\epsilon_i)dk$, and $K_{ss} = \sum_i \int k\bar{\mu}(\epsilon_i, k)dk = \sum_i K_{ss}(\epsilon_i)\phi(\epsilon_i)$ ($\bar{\mu}(\epsilon_i, k)$ is the stationary distribution of (k, ϵ_i)).[14]

Bias correction When we allow for idiosyncratic shocks, we perform bias correction for each level of idiosyncratic shocks. Let $K_{ss}(\epsilon_i) = \int \bar{g}_k(\epsilon_i, k)\bar{\mu}_k(k|\epsilon_i)dk$ be the amount of ϵ_i–indexed capital in the steady state, which is obtained by the decision rule $\bar{g}_k(\epsilon_i, k)$ and the stationary distribution $\bar{\mu}(\epsilon_i, k)$. Also let $\tilde{K}_{ss}(\epsilon_i) = \bar{g}_k(\epsilon_i, \int k\bar{\mu}_k(k|\epsilon_i)dk) = \bar{g}_k(\epsilon_i, K_{ss}(\epsilon_i))$ as the decision rule evaluated at $K_{ss}(\epsilon_i)$. Then the bias–correction term depends on each ϵ_i, $\zeta_k(\epsilon_i) \equiv K_{ss}(\epsilon_i) - \tilde{K}_{ss}(\epsilon_i) = K_{ss}(\epsilon_i) - \bar{g}_k(\epsilon_i, K_{ss}(\epsilon_i))$, and we have

$$K_{t+1} = \sum_i \left[g_k(\epsilon_i, K_t(\epsilon_i); z_t, K_t) + \zeta_k(\epsilon_i)\right]\phi(\epsilon_i).$$

Discussion Terry (2017) and Sunakawa (2012) undertake a naive explicit aggregation. That is, they evaluate the policy function $g_k(\epsilon_i, k; z_t, K_t)$ at $k = K_t$ instead of $k = K_t(\epsilon_i)$. However, this aggregation may lead to large forecasting errors, as shown in Terry (2017). den Haan and Rendahl (2010) assume that the individual decision rule is a function of each ϵ–indexed capital. For example, in the standard Krusell and Smith (1998) model, the idiosyncratic shock takes two values, $\epsilon_i \in \{u, e\}$, and the policy function becomes $g(\epsilon_i, k; z_t, K_t(u), K_t(e))$. We evaluate this at $k = K_t(\epsilon_i)$ for each ϵ_i as we know the value of ϵ–indexed capital as a grid point. This approach is very costly or infeasible as the number of grid points for ϵ_i increases because we need more dimensions of the state space.

We use steady–state information for both the bias correction and the ϵ–indexed aggregation. Therefore, an accurate steady–state solution is necessary. We find that non–stochastic methods (Young, 2010) augmented with the use of eigenvalue and eigenvector decomposition often yield a precise approximation to the stationary distribution. See Appendix A.1 for details. In contrast, stochastic methods may give a poor approximation to the stationary equilibrium that also worsens

[14] The assumption that $\psi(\epsilon_i) = K_t(\epsilon_i)/K_t$ is time–invariant may not be valid with large aggregate shocks when the share of each ϵ–indexed capital significantly changes over time. Later we relax this assumption with an exogenously time–varying employment measure in Section 5.

the precision of the Xpa algorithm.[15]

4 Numerical examples

In this section, we present the numerical results from applying the KS and Xpa algorithms to the heterogeneous–firm models in KT (2003; 2008). We examine both cases without idiosyncratic shocks (KT 2003, aka the traditional model) and with idiosyncratic shocks (KT 2008, aka the extended model), although we mainly focus on the latter in the main text.

4.1 Parameterization

The parameter values used in this paper closely follow those in KT (2003; 2008). The functional forms are $U(C, 1-N) = \log C - \eta N$ and $F(k, n) = k^{\alpha} n^{\nu}$, where $\eta > 0$, $\alpha \in (0, 1)$, $\nu \in (0, 1)$, and $\alpha + \nu < 1$. The distribution of the random fixed costs drawn by firms is uniform $G : [0, \bar{\xi}] \to [0, 1]$. The Markov chain for aggregate– and firm–level productivity approximates an AR(1) process by Tauchen's (1986) method; $z' = \rho_z z + \eta'_z$ and $\varepsilon' = \rho_\varepsilon \varepsilon + \eta'_\varepsilon$ for each, where $\eta'_z \sim N(0, \sigma_{z\eta})$ and $\eta'_\varepsilon \sim N(0, \sigma_{\varepsilon\eta})$. The range of constrained investment is given by $\Omega(k) = \left[\frac{1-\delta+a}{\gamma}k, \frac{1-\delta+b}{\gamma}k\right]$, where $-a = b > 0$.

The parameter values and targeted moments are in Tables 1 and 2, respectively. The unit of time is a year. The mean growth rate of technology γ implies a 1.6 percent average annual growth rate. The discount factor β is from an average annual real interest rate $\beta\gamma^{-1} - 1$ of 4 percent. The annual depreciation rate is 6.9 percent to match the average investment–capital ratio in the fixed asset tables. Labor's share ν is 0.64 as in Prescott (1986).

We internally calibrate the other parameters in the model. The capital share α is set to match a capital–output ratio of 2.353. η is a normalization parameter and set so that the steady state labor is 1/3. The values of $(\bar{\xi}, b, \sigma_{\xi\eta})$ are chosen to match the data of the average firm–level investment rates in Cooper and Haltiwanger (2006) as shown in Table 2. Our steady–state results show that we match these internal calibration targets quite well. In the traditional model of KT (2003), firm–

[15] Rios-Rull (1999) mentions that non–stochastic methods provide a better approximation of the steady state than stochastic methods.

level productivity is constant and there is no constrained investment, i.e., $b = \sigma_{\xi\eta} = 0$. $\bar{\xi} = 0.014$ is set to match the average investment rate of positive spikes ($i/k > 0.20$). In the extended model of KT (2008), the other moments are reasonably matched between the model and the data. Note that we set $\rho_\varepsilon = \rho_z$ as there is little agreement about the persistence of the firm–level idiosyncratic productivity.

Table 1: Parameters.

a. Externally calibrated

	β	γ	δ	ν	ρ_z	$\sigma_{z\eta}$	ρ_ε
Traditional	0.977	1.016	0.069	0.640	0.859	0.014	—
Extended	0.977	1.016	0.069	0.640	0.859	0.014	0.859

b. Internally calibrated

	α	η	$\bar{\xi}$	b	$\sigma_{\varepsilon\eta}$
Traditional	0.256	2.400	0.014	0.0	—
Extended	0.256	2.400	0.0083	0.011	0.022

Notes: The values of $(\xi, b, \sigma_{\varepsilon\eta})$ are chosen by following KT (2008).

Table 2: Targeted moments.

a. Aggregate variables (without aggregate uncertainty)

	K/Y	N
Traditional	2.3487	0.3337
Extended	2.3515	0.3338

b. Firm–level investment rates (without aggregate uncertainty)

	Mean	Std.dev.	Inaction	Spike+	Spike-	Invest+	Invest-
Traditional	0.1046	0.0214	0.2113	0.1284	0.0000	0.7887	0.0000
Extended	0.1158	0.0716	0.0097	0.1595	0.0130	0.8127	0.1779
Census data	0.122	—	0.081	0.186	0.018	0.815	0.104

Notes: Census data are from Cooper and Haltiwanger (2006). Inaction: $|i/k| < 0.01$, Spike+: $i/k \geq 0.2$, Spike-: $i/k \leq -0.2$, Invest+: $i/k \geq 0.01$, and Invest-: $i/k \leq -0.01$.

4.2 Model statistics

To start, we detail the unconditional business cycle statistics in Table 3. We simulate the model for 2,500 periods and discard the first 500 periods. We take logs of the simulated data and use the Hodrick–Prescott filter to remove the low frequency components. We can then see that after being detrended, the outcomes of the two algorithms are very similar. This is also true for the higher–order moments, including skewness and excess kurtosis, of the aggregate investment–capital ratio, although estimates of these higher–order moments are usually unstable.

As we can see from the plot of the simulated time series of the aggregate variables in Figure 2, not only the second– and higher–order moments, but also the levels are very close to each other for the two algorithms. This is through the bias–correction procedure explained in Section 3. Compared with previous studies that apply the Xpa algorithm (Terry, 2017; Sunakawa, 2012), the ϵ–indexed aggregation with bias correction yields almost indistinguishable results in both levels and higher–order moments from those with the KS algorithm.

We also consider the conditional moments by computing the impulse responses to a positive shock to the aggregate productivity. We simulate the economy with no shocks for a long period so that the economy reaches the stochastic steady state (Coeurdacier et al., 2011). Then we hit the economy with a one–time persistent positive productivity shock. As depicted in Figure 3, we can see that the outcomes of the two algorithms are again quite similar, although the responses of aggregate capital as computed by the Xpa algorithm are slightly weaker.

Table 3: Business cycle statistics, extended model.

	Y	C	I	N	K	Z
Standard deviation						
KS	(2.5329)	0.4344	3.6109	0.6652	0.5138	0.5972
Xpa	(2.5100)	0.4401	3.5767	0.6564	0.5097	0.6027
Output correlation						
KS	1.0000	0.8732	0.9731	0.9364	0.0459	0.9996
Xpa	1.0000	0.8802	0.9734	0.9360	0.0449	0.9995

	Persist.	Std.dev.	Skewness	Exc.kur.
Agg. investment rate, I/K				
KS	0.6499	0.0107	0.0714	-0.0943
Xpa	0.6525	0.0105	0.0575	-0.0915

Notes: We take logs of the simulated data and use the Hodrick–Prescott filter (with a filtering parameter $\lambda = 100$) to remove the low–frequency components.

Figure 2: Unconditional business cycle simulation, extended model.

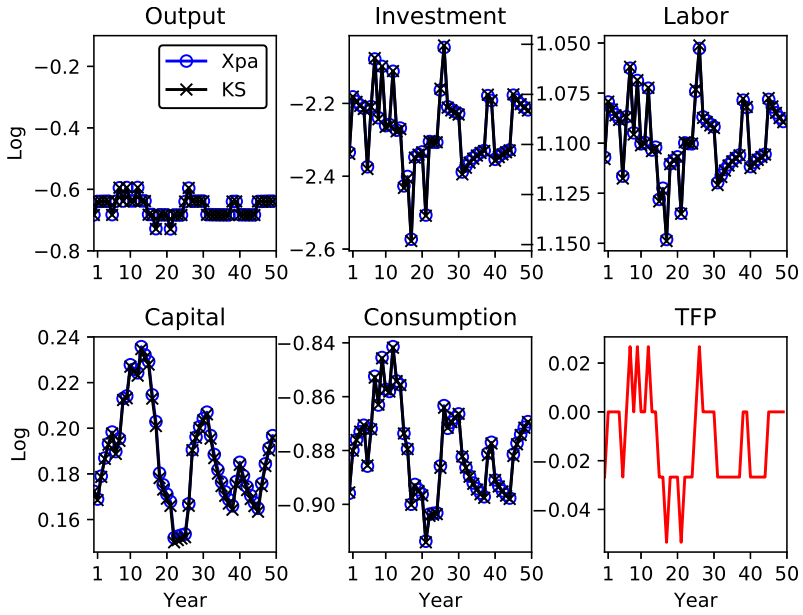


Figure 3: Impulse responses, extended model.

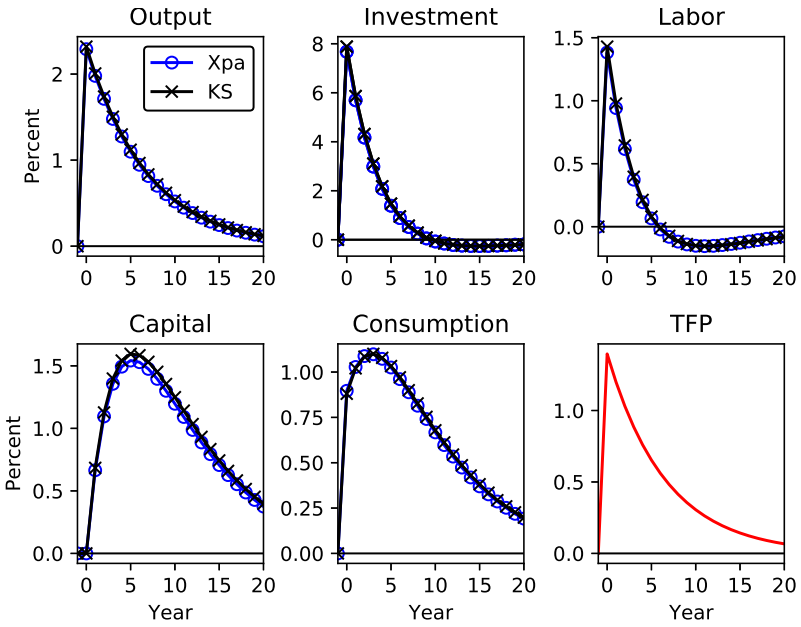


Table 4: Level differences between KS and Xpa, extended model.

a. Aggregate variables

	K/Y	N	$p = C^{-1}$
KS	2.3506	0.3338	2.3666
Xpa	2.3538	0.3339	2.3656
SS	2.3515	0.3338	2.3663

b. Firm–level investment rates

	Mean	Std.dev.	Inaction	Spike+	Spike-	Invest+	Invest-
KS	0.1158	0.0717	0.0098	0.1595	0.0130	0.8126	0.1779
Xpa	0.1158	0.0717	0.0097	0.1591	0.0130	0.8126	0.1779
SS	0.1158	0.0716	0.0097	0.1595	0.0130	0.8127	0.1779

Notes: Inaction: $|i/k| < 0.01$, Spike+: $i/k \geq 0.2$, Spike-: $i/k \leq -0.2$, Invest+: $i/k \geq 0.01$, and Invest-: $i/k \leq -0.01$.

Table 4 displays the aggregate and disaggregate variables in levels at the stochastic steady state. Interestingly, the levels of the capital–output ratio and the hours worked are almost the same as

in the deterministic steady state. This is also true when we look at the distribution of firm–level investment rates. This means that the higher–order effects of aggregate uncertainty on the levels of aggregate and disaggregate variables are very limited, as opposed to the heterogeneous–household models of Krusell and Smith (1998) in which households are risk averse.

In the traditional KT model (2003), we also observe a strong similarity in the unconditional and conditional moments between the outcomes of the Xpa and KS algorithms. The corresponding tables and figures are in Appendix A.3.

4.3 Accuracy and efficiency

To evaluate accuracy in terms of forecasting errors, we employ not only the R–squared and the root mean squared errors (RMSE), but also the den Haan (DH) statistics based on the forecasting rules $\Gamma_K(K_t, z_t)$ and $\Gamma_p(K_t, z_t)$ as in den Haan (2010). We calculate both the static and dynamic forecast errors as in Terry (2017). To calculate the dynamic forecast errors, taking as given the sequence of exogenous shock $\{z_t\}_{t=1}^T$ and the initial value of K_1, we use the forecasting rules recursively to generate the sequence of the next period's aggregate capital and the shadow price, $\{\tilde{K}_{t+1}, \tilde{p}_t\}_{t=1}^T$. The sequence is compared with the simulated data from the model, $\{K_{t+1}, p_t\}$. Then we have

$$u_{K,t} = \left| \tilde{K}_{t+1} - K_{t+1} \right|,$$
$$u_{p,t} = \left| \tilde{p}_t - p_t \right|,$$

for $t = 1, ..., T$. The static forecast errors are one–step ahead forecast errors, i.e., $e_{K,t} = |\Gamma_K(K_t, z_t) - K_{t+1}|$ and $e_{p,t} = |\Gamma_p(K_t, z_t) - p_t|$. We sort the forecast errors by the realization of z_t for each z_i.[16] The dynamic forecast errors are more robust measures of accuracy as errors accumulate over time, and we report the max and mean DH errors based on the dynamic forecast errors.

Figure 4 is the so–called den Haan fundamental plot for aggregate capital and consumption (the inverse of the shadow price). We compare the simulated data with the dynamic and static

[16] Using the static forecast errors sorted for each z_i with T_i samples, we also have RMSEs as $\sqrt{\sum_{t=1}^{T_i} e_{K,t}^2}$ and $\sqrt{\sum_{t=1}^{T_i} e_{p,t}^2}$ and R–squares as $1 - \sum_{t=1}^{T_i} e_{K,t}^2 / \sum_{t=1}^{T_i} (K_{t+1} - \bar{K})^2$ and $1 - \sum_{t=1}^{T_i} e_{p,t}^2 / \sum_{t=1}^{T_i} (p_t - \bar{p})^2$ where $\bar{K} = (1/T_i) \sum_{t=1}^{T_i} K_{t+1}$ and $\bar{p} = (1/T_i) \sum_{t=1}^{T_i} p_t$.

forecasts. We can see that the differences between the simulated data and the dynamic or static forecasts (i.e., u_{Kt} and u_{pt} or e_{Kt} and e_{pt}) are very small throughout. Specifically, by using the extended model, we report in Table 5 that the maximum values of the DH statistics in the Xpa algorithm are at most twice those in the KS algorithm (0.792 vs. 0.337 for u_{Kt} and 0.263 vs. 0.119 for u_{pt}). Indeed, they are much smaller (by a factor of three or four times) than what Terry (2017) reports for the max DH statistics (3.457 for u_{Kt} and 1.210 for u_{pt}). In the traditional model, the difference is even smaller as the Xpa algorithm is almost equally as accurate as the KS algorithm.

Figure 4: den Haan fundamental plots, extended model.

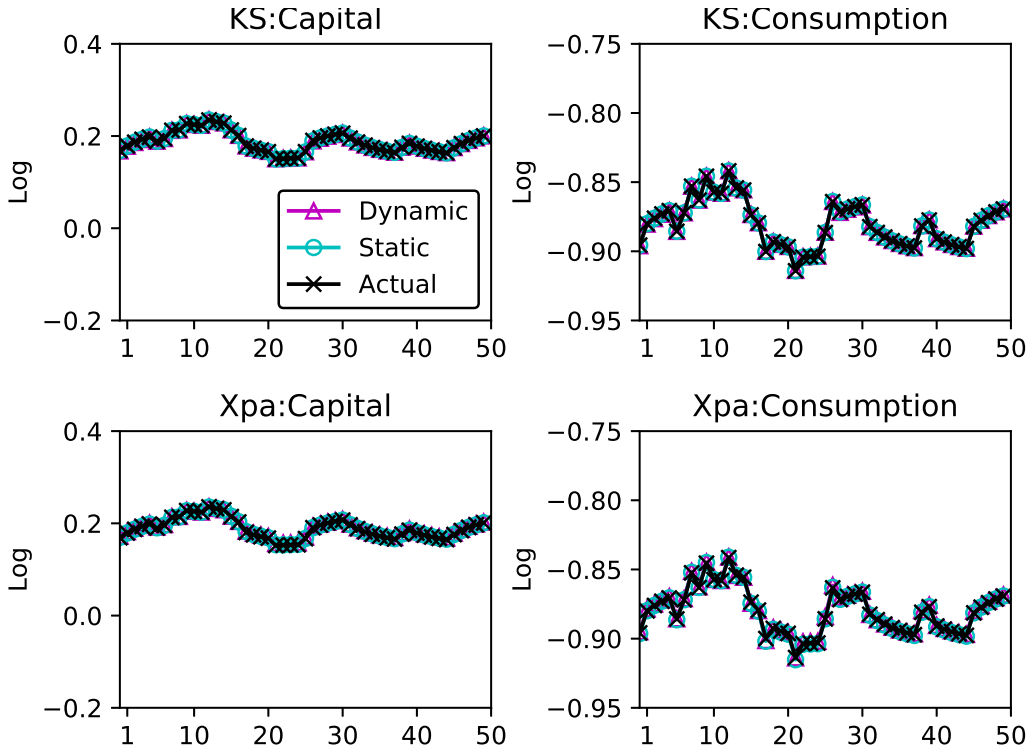


Table 5: DH statistics, RMSE, and R–squared.

a. Extended model

		DHmax, %		DHmean, %		RMSE, %		R–squared	
		K'	p	K'	p	K'	p	K'	p
KS	z_1	0.3095	0.1174	0.0055	0.0021	0.0256	0.0064	0.9999	1.0000
	z_2	0.3386	0.1189	0.0166	0.0063	0.0274	0.0060	0.9999	1.0000
	z_3	0.2605	0.1122	0.0178	0.0063	0.0270	0.0056	0.9999	1.0000
	z_4	0.3064	0.1062	0.0149	0.0052	0.0256	0.0064	0.9999	1.0000
	z_5	0.2297	0.0743	0.0040	0.0016	0.0184	0.0056	0.9999	1.0000
Xpa	z_1	0.7921	0.1858	0.0404	0.0055	0.1660	0.1574	0.9966	0.9873
	z_2	0.6629	0.2535	0.0831	0.0197	0.0820	0.0701	0.9993	0.9978
	z_3	0.5470	0.2631	0.0558	0.0228	0.0379	0.0243	0.9999	0.9997
	z_4	0.3941	0.2341	0.0165	0.0145	0.0196	0.0312	1.0000	0.9995
	z_5	0.2227	0.1892	0.0032	0.0057	0.0243	0.0600	0.9999	0.9973

b. Traditional model

		DHmax, %		DHmean, %		RMSE, %		R–squared	
		K'	p	K'	p	K'	p	K'	p
KS	z_1	0.3255	0.1490	0.0055	0.0023	0.0293	0.0189	0.9999	0.9998
	z_2	0.3675	0.1528	0.0181	0.0075	0.0314	0.0184	0.9999	0.9999
	z_3	0.2945	0.1407	0.0197	0.0080	0.0307	0.0170	0.9999	0.9999
	z_4	0.3389	0.1271	0.0159	0.0062	0.0276	0.0160	0.9999	0.9999
	z_5	0.2371	0.0866	0.0042	0.0019	0.0195	0.0134	0.9999	0.9999
Xpa	z_1	0.4755	0.1374	0.0256	0.0046	0.0941	0.0708	0.9989	0.9975
	z_2	0.4457	0.1624	0.0628	0.0158	0.0620	0.0364	0.9996	0.9994
	z_3	0.4125	0.1769	0.0559	0.0214	0.0366	0.0171	0.9999	0.9999
	z_4	0.3425	0.1779	0.0277	0.0172	0.0216	0.0268	0.9999	0.9997
	z_5	0.2747	0.1706	0.0072	0.0063	0.0149	0.0405	1.0000	0.9988

Notes: Taking the dynamic and static forecast errors (sorted for each z_i) $\{u_{y,t}\}_{t=1}^{T_i}$ and $\{e_{y,t}\}_{t=1}^{T_i}$ for $y \in \{K, p\}$ as given, DHmax is the maximum value of $\{u_{y,t}\}_{t=1}^{T_i}$, DHmean is the average of $\{u_{y,t}\}_{t=1}^{T_i}$, RMSE is $\sqrt{\sum_{t=1}^{T_i} e_{y,t}^2}$, and R–squared is $1 - \sum_{t=1}^{T_i} e_{y,t}^2 / \sum_{t=1}^{T_i} (y_t - \bar{y})^2$ where $\bar{y} = (1/T) \sum_{t=1}^{T_i} y_t$.

Finally, in Table 6, we compare the two algorithms in terms of computation time. The programming code is written in Fortran with OpenMP directives so that all the cores on a computer are utilized. We find the Xpa algorithm is about 20–100 times faster than the KS algorithm. This result is comparable with that obtained by den Haan and Rendahl (2010) for the Krusell and Smith (1998) model. Specifically, the speed gain is mainly in the outer loop to obtain the forecasting rules,

as we do not need to simulate the model to update the forecasting rules in the Xpa algorithm. Note that the two algorithms share the same subroutines for calculating the steady state, inner loops, and unconditional/conditional simulations.

Table 6: Computational time.

	Inner	Outer	# of loop	Total
Extended				
KS	0.750	122.607	10.000	1233.570
Xpa	0.745	0.032	11.000	8.547
Xpa/KS	0.993	<0.001	1.100	0.007
Traditional				
KS	0.386	6.532	6.000	41.505
Xpa	0.386	0.007	6.000	2.355
Xpa/KS	1.000	0.001	1.000	0.057

Notes: Computation using a workstation with Xeon E5–2696v4 (2.2 Ghz, 44 cores).

5 Krueger–Mitman–Perri models

Krueger, Mitman and Perri (2016; hereafter KMP) extend the analysis in Krusell and Smith (1998) and show that micro–level heterogeneity matters for the aggregate dynamics.[17] In this section, we show that the Xpa algorithm can be applied to the models in KMP.

5.1 Model

In the economy, there are the representative firm, heterogeneous households, and the government. The continuum of households differs in their asset holdings, denoted as k, and the other idiosyncratic shocks, represented by a vector $x_i \equiv (e_i, y_i, \beta_i)$, which include the following.

- Employment status e_i: $e_i \in \{1, \rho\}$, where ρ is the replacement ratio. That is, households have unemployment insurance.

[17] We use the model in a working paper version of Krueger et al. (2015). There are three major differences from Krueger et al. (2016): (i) earnings risk is an AR(1) process instead of a joint process combining persistent and transitory components, (ii) there is no retirement and social insurance, which influences (iii) the calibration of β. Nonetheless, introducing retirement and social insurance and/or a more complicated stochastic process for earnings risk is relatively straight–forward.

- Labor productivity y_i: Households face idiosyncratic earnings risk,

$$\log y' = \phi \log y + \eta',$$

  where $\eta' \sim N(0, \sigma_\eta^2)$. We transform the continuous variables in the AR(1) process into the discretized grid points of a Markov chain. That is, $y_i \in \{y_1, ..., y_{N_y}\}$.

- Discount factor β_i: Households are born with $\beta_i \in [\bar{\beta} - \epsilon, \bar{\beta} + \epsilon]$ and live with β their entire life. They die with probability $(1 - \theta) \in [0, 1]$. A fraction θ of households is born in every period with no assets.

A household is defined by the individual state variables (x_i, k). The aggregate state of the economy is given by (z_t, μ_t), where z_t is aggregate productivity and μ_t is the distribution of individual assets and the other characteristics held by each household. Taking as given the wage rate w_t and the real interest rate r_t, the household (x_i, k) chooses consumption c and the next period's capital k'. The household's life–time utility from choosing (c, k') is

$$\begin{aligned} v(x_i, k; z_t, \mu_t) &= \max_{c, k' \geq 0} \{u(c) \\ &+ \theta\beta_i \sum_{z_{t+1}} \sum_{e'} \sum_{y'} \pi_z(z_{t+1}|z_t)\pi_e(e'|e_i, z_{t+1}, z_t)\pi_y(y'|y_i) \\ &\quad \times v(e', y', \beta_i, k'; z_{t+1}, \mu_{t+1})\big\}, \end{aligned} \tag{12}$$

subject to the budget constraint and forecasting rule for the aggregate state in the next period

$$\begin{aligned} c + k' &= (1 - \tau(z_t; \rho))w_t y_i e_i + (1 + r_t - \delta)k/\theta, \\ \mu_{t+1} &= \Gamma(z_t, \mu_t), \end{aligned}$$

where each of $\pi_z(z_{t+1}|z_t)$, $\pi_e(e'|e_i, z_{t+1}, z_t)$, and $\pi_y(y'|y_i)$ is the conditional probability of the next period's aggregate or idiosyncratic state given the current period's state. $\tau(z_t; \rho)$ is a flat income tax with which the government finances unemployment insurance. By solving the household's problem,

we have the individual decision rule for saving

$$k' = g(x_i, k; z_t, \mu_t), \tag{13}$$

which is aggregated as

$$K_{t+1} = \sum_i \int g(x_i, k; z_t, \mu_t)\theta\mu_t(x_i, k)dk,$$

where $\mu_t(x_i, k)$ is the measure of households with the amount of individual asset k and other peculiarities x_i and θ is the survival probability from period t to period $t+1$. Note that the replacing newborns have no assets.

The representative firm has access to a Cobb–Douglas production technology $Y_t = z_t K_t^{\alpha} L_t^{1-\alpha}$ and maximizes profit $Y_t - (r+\delta)K_t - w_t L_t$ in each period. The first–order necessary conditions are

$$w_t = (1-\alpha) z_t K_t^{\alpha} L_t^{-\alpha},$$
$$r_t = \alpha z_t K_t^{\alpha-1} L_t^{1-\alpha} - \delta,$$

where $K_t = \sum_i \int k\mu_t(x_i, k)dk$ is the aggregate capital and $L_t = \sum_i \int \mathbb{I}(e_i = 1)\mu_t(k, x_i)dk$ is the amount of aggregate hours worked.

A recursive competitive equilibrium is defined so as to satisfy (i) each household's optimality, (ii) the firm's optimality, (iii) feasibility: $C_t = Y_t + (1-\delta)K_t - K_{t+1}$, and (iv) consistency between the individual decision rules and the aggregate forecasting rules.

ϵ–indexed aggregation with time–variant employment measure We assume that approximate aggregation holds. That is, $v(x_i, k; z_t, \mu_t)$ and $g(x_i, k; z_t, \mu_t)$ are replaced by $v(x_i, k; z_t, K_t)$ and $g(x_i, k; z_t, K_t)$ and the aggregate forecasting rule is given by $K_{t+1} = \Gamma(z_t, K_t)$. The aggregate capital in the next period is given by

$$K_{t+1} = \sum_i \int g(x_i, k; z_t, K_t)\theta\mu_t(x_i, k)dk,$$

The measure of (x_i, k) can be decomposed into

$$\mu_t(x_i, k) = \mu_{kt}(k|x_i)\phi(x_i; z_t),$$

where $\phi(x_i; z_t) = \int \mu_t(x_i, k)dk$ is the marginal distribution of x_i that is assumed to be a *time–variant* measure of x_i depending on z_t.[18] This implies exogenously time–varying hours worked

$$\begin{aligned} L_t &= \sum_i \int \mathbb{I}(e_i = 1)\mu_t(k, x_i)dk \\ &= \sum_i \mathbb{I}(e_i = 1)\phi(x_i; z_t) \\ &= 1 - u(z_t) \end{aligned}$$

With the time–varying employment measure $\phi(x_i; z_t)$, the aggregation in the Xpa algorithm becomes

$$\begin{aligned} K_{t+1} &= \sum_i \int g(x_i, k; z_t, K_t)\theta\mu_{kt}(k|x_i)\phi(x_i; z_t)dk, \\ &\approx \sum_i g(x_i, \int k\mu_{kt}(k|x_i)dk; z_t, K_t)\theta\phi(x_i; z_t), \\ &= \sum_i g(x_i, K_t(x_i; z_t); z_t, K_t)\theta\phi(x_i; z_t), \end{aligned}$$

where $K_t(x_i; z_t) \equiv \int k\mu_{kt}(k|x_i)dk$ is the amount of capital indexed by x_i at the beginning of period t. As before, $K_t(x_i; z_t)$ is not on the grid points and we evaluate the individual decision rule at $K_t(x_i; z_t) = \psi(x_i; z_t)K_t$ where $\psi(x_i; z_t) = K_{ss}(x_i; z_t)/K_{ss}$ and $K_{ss}(x_i; z_t) = \int k\bar{\mu}_k(k|x_i; z_t)dk$.

We also calculate the bias–correction terms

$$\zeta(x_i) = K_{ss}(x_i) - \bar{g}(K_{ss}(x_i)),$$

where $K_{ss}(x_i) = \int \bar{g}_k(\epsilon_i, k)\bar{\mu}_k(k|x_i)dk$. Note that the bias–correction terms do not depend on z_t.[19]

[18] $\mu_{kt}(k|x_i) = \frac{\mu_t(x_i,k)}{\int \mu_t(x_i,k)dk}$ is the conditional distribution of $\mu_t(x_i, k)$, which is also time–variant.

[19] There are two different decompositions of the stationary distribution here. One is $\bar{\mu}(x_i, k) = \bar{\mu}_k(k|x_i)\phi(x_i)$ and

Then we have

$$K_{t+1} = \sum_i \left[g(x_i, K_t(x_i; z_t); z_t, K_t)\theta + \zeta(x_i) \right] \phi(x_i; z_t).$$

5.2 Parameterization

We calibrate the model to quarterly US data, where the capital share is $\alpha = 0.36$ and the depreciation rate is $\delta = 0.025$. A two–state Markov process is used for aggregate productivity with transition matrix

$$\pi_z(z'|z) = \begin{pmatrix} \rho_l & 1-\rho_l \\ 1-\rho_h & \rho_h \end{pmatrix}.$$

We have two calibration targets to determine the persistence parameters (p_l, p_h): the average length of a severe recession $EL_l = 1/(1-\rho_l)$ and the fraction of time in a severe recession $\Pi_l = (1-\rho_h)/(2-\rho_l-\rho_h)$. With $EL_l = 22$ quarters and $\Pi_l = 16.48\%$ obtained from the sample periods, we have $(\rho_l, \rho_h) = (0.9545, 0.9910)$.[20]

To identify the levels of aggregate technology at each state, we also target $Y_l/Y_h = 0.9298$, corresponding to a fall in GDP per capita during the great recession of about 7%. We also have the average unemployment rates $u(z_l) = 8.39\%$ and $u(z_h) = 5.33\%$ from the sample periods. Then we have $z_l/z_h = (Y_l/L_l)/(Y_h/L_h) = 0.9608$ assuming $L_l/L_h = K_l/K_h$. A normalization $\Pi_l z_l + (1-\Pi_l) z_h = 1$ provides $(z_l, z_h) = (0.9676, 1.0064)$. For the employment status, Krueger et al. (2016) use the job finding and separation rates from CPS and obtain transition matrices

the other is $\bar{\mu}(x_i, k) = \bar{\mu}_k(k|x_i; z_t)\phi(x_i; z_t)$.

[20] Krueger et al. (2016) define a recession to be one where the unemployment rate rises above 9% for one quarter and continues to be above 7%. During 1948.I–2014.III, two recession periods, 1980.II–1986.II and 2009.I–2013.III are duly identified.

depending on the current and next period's aggregate technology $z, z' \in \{z_l, z_h\}$:

$$\pi_e(e'|e, z_l, z_l) = \begin{pmatrix} 0.3378 & 0.6622 \\ 0.0606 & 0.9394 \end{pmatrix}, \quad \pi_e(e'|e, z_h, z_h) = \begin{pmatrix} 0.1890 & 0.8110 \\ 0.0457 & 0.9543 \end{pmatrix},$$
$$\pi_e(e'|e, z_h, z_l) = \begin{pmatrix} 0.3382 & 0.6618 \\ 0.0696 & 0.9304 \end{pmatrix}, \quad \pi_e(e'|e, z_l, z_h) = \begin{pmatrix} 0.2220 & 0.7780 \\ 0.0378 & 0.9622 \end{pmatrix}.$$

Note that these matrices are consistent with the time–varying unemployment rates $u(z_l)$ and $u(z_h)$.[21] For the labor productivity, Krueger et al. (2016) estimate $(\hat{\phi}, \hat{\sigma}_y^2)$ from the annual PSID data and translate them for the quarterly model, $(\phi, \sigma_y^2) = (0.9457, 0.0359)$. Rouwenhorst's (1995) method instead of Tauchen (1986) is used for discretization because of the long persistence ϕ. The set of discount factors $B = [\bar{\beta} - \epsilon, \bar{\beta} + \epsilon]$ is divided into grid points where $(\bar{\beta}, \epsilon) = (0.98349, 0.01004)$. $\theta = 0.99375$ so that the probability of death is $1/160$. For the replacement ratio, $\rho = 0.5$ is used for the benchmark case. Note that $\tau(z_t; \rho) = \frac{u(z_t)\rho}{1-u(z_t)+u(z_t)\rho}$ holds from the government's balanced budget. The parameter descriptions and values are summarized in Table 7.

[21] The stationary distribution of employment status implied by $\pi_e(e'|e, z_l, z_l)$ ($\pi_e(e'|e, z_h, z_h)$) coincides with the employment rate $1 - u(z_l)$ ($1 - u(z_h)$).

Table 7: Parameters.

Parameters		
Capital share	α	0.36
Depreciation rate	δ	0.025
Risk aversion	σ	1.0
Agg. productivity	(z_l, z_h)	$(0.9676, 1.0064)$
Unemployment rate	(u_l, u_h)	$(0.0839, 0.0533)$
Transition for agg. prod.	$\pi_z(z'\|z)$	$\begin{pmatrix} 0.9545 & 0.0455 \\ 0.0090 & 0.9910 \end{pmatrix}$
Replacement ratio	ρ	0.50
Transition for emp. status	$\pi_e(e'\|e, z_h, z_h)$	$\begin{pmatrix} 0.1890 & 0.8110 \\ 0.0457 & 0.9543 \end{pmatrix}$
	$\pi_e(e'\|e, z_h, z_l)$	$\begin{pmatrix} 0.3382 & 0.6618 \\ 0.0696 & 0.9304 \end{pmatrix}$
	$\pi_e(e'\|e, z_l, z_h)$	$\begin{pmatrix} 0.2220 & 0.7780 \\ 0.0378 & 0.9622 \end{pmatrix}$
	$\pi_e(e'\|e, z_l, z_l)$	$\begin{pmatrix} 0.3378 & 0.6622 \\ 0.0606 & 0.9394 \end{pmatrix}$
Persistence of idio. prod.	ϕ	0.9457
Std. dev. of idio. prod.	σ_y^2	0.0359
Discount factor	$\bar{\beta}$	0.98349
Discount factor shock	ϵ	0.01004
Survival rate	θ	0.99375

Source: Krueger, Mitman, and Perri (2015, working paper).

In Table 8, we replicate the stationary distribution results in KMP. We consider three calibrations. The baseline calibration is denoted by KMP. In the KS calibration, we omit the earnings risk and discount factor heterogeneity. We also set $\theta = 1$ and $\rho = 0.01$. That is, we have no stochastic death and unemployment insurance (we need a positive $\rho > 0$ so that the amount of consumption is positive). We add the earnings risk only to the KS calibration, denoted $+\sigma(y)$. The KMP calibration shows that there are many households that have no net wealth. In contrast, the top 20% of households hold most wealth (82.7% in the 2006 PSID vs. 82.9% in the model).[22] The model with the baseline calibration also has a Gini coefficient of 0.78, which is comparable with the data.

The KS calibration has low wealth inequality and a small Gini coefficient (0.32). In our com-

[22] Unfortunately, the model cannot generate the concentration in the top 1% as observed in the data.

putation, the $+\sigma(y)$ calibration lies in between the KMP and KS calibrations in terms of wealth inequality. The three calibrations have a similar capital–output ratio at the aggregate level, although the KS calibration has a slightly low value.

Table 8: Net wealth distribution (without aggregate uncertainty).

	Data		Models		
% of Share:	PSID, 06	SCF, 07	KMP	KS	$+\sigma(y)$
Q1	-0.9	-0.2	0.1182	7.3694	1.7619
Q2	0.8	1.2	0.8796	12.3381	6.4322
Q3	4.4	4.6	3.5852	17.1474	13.6520
Q4	13.0	11.9	12.5377	23.7850	24.9265
Q5	82.7	82.5	82.8793	39.3601	53.2275
90–95	13.7	11.1	19.2195	9.9165	13.4301
95–99	22.8	25.3	29.3285	9.9975	14.5831
T1	30.9	33.5	16.0137	3.4700	5.4405
Wealth Gini	0.77	0.78	0.7843	0.3205	0.5162
K/Y	—	—	11.0821	10.2674	11.5596

Notes: The figures in columns "Data" (2006 PSID and 2007 SCF) are from Krueger et al. (2016). KS: Remove the earnings risk and discount factor heterogeneity and set $\theta = 1$ and $\mu = 0.01$. $+\sigma(y)$: Remove the discount factor heterogeneity and set $\theta = 1$ and $\rho = 0.01$.

5.3 Model statistics

To start, we compute the unconditional business cycle statistics for each calibration in Table 9. We can see that the two algorithms yield very similar results in all three specifications. The outcomes are also similar in the second–order moments as well as in the levels. Specifically, we have a high correlation between output and consumption, volatile investment, and persistent output. The baseline KMP calibration has the largest procyclicality of consumption, less volatile investment, and less persistent output. Figure 5 also depicts that the unconditional simulation paths for each algorithm are close to each other in the KMP calibration. This is also the case for the other two calibrations. See Appendix B.3.

Table 9: Business cycle statistics.

Model	Mean(K/Y)	Corr(Y,C)	Std.dev.(I)	Corr(Y,Y_{-4})
KMP				
KS	11.0627	0.9405	0.0490	0.8851
Xpa	11.0624	0.9391	0.0494	0.8856
KS				
KS	10.2887	0.9208	0.0517	0.8947
Xpa	10.2873	0.9226	0.0512	0.8940
$+\sigma(y)$				
KS	11.4663	0.9139	0.0567	0.8911
Xpa	11.4644	0.9168	0.0567	0.8913

Notes: All variables unfiltered and in levels as in Krusell and Smith (1998).

Figure 5: Unconditional business cycle simulation, KMP calibration.

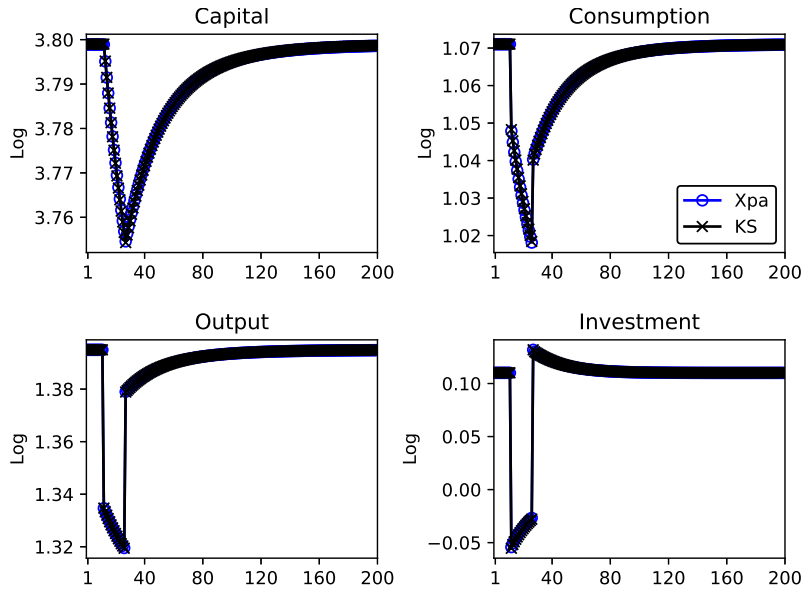


Table 10 provides the micro–level moments at the stochastic steady state as well as the aggregate capital–output ratio and the fall in aggregate consumption from a negative aggregate productivity shock. Once again, the results from the two algorithms are almost indistinguishable. In both algorithms, the capital–output ratio is slightly higher than in the deterministic steady state because of precautionary saving in the presence of aggregate uncertainty (see also Table 8). This cannot be captured by Reiter's (2009) linearization method. The consumption decline in response to a one–

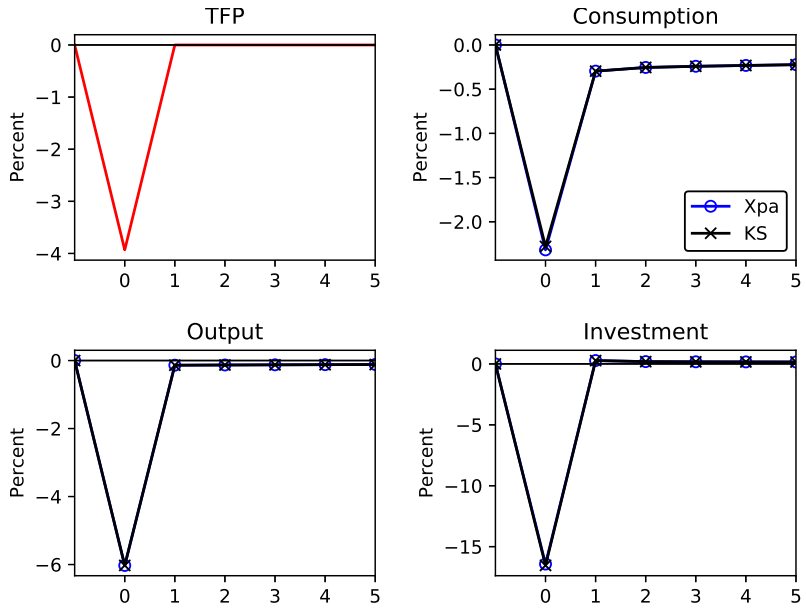
time negative shock is also similar. We confirm this in Figure 6 by plotting the impulse responses for the KMP calibration in which heterogeneity matters.

Interestingly, in the KS calibration, the micro–level moments with aggregate uncertainty are slightly different from those in the stationary distribution. This is also because of precautionary saving. With aggregate uncertainty, the poor tend to save more, which makes the rich dissave as the interest rate falls (K/Y becomes higher compared with the deterministic steady state). As a result, the wealth Gini is even smaller in the stochastic steady state than in the deterministic steady state.

Table 10: Net wealth distribution (with aggregate uncertainty) and consumption decline.

	Models					
	KMP		KS		$+\sigma(y)$	
% of Share:	KS	Xpa	KS	Xpa	KS	Xpa
Q1	0.1159	0.1158	8.3503	8.3017	1.7469	1.7467
Q2	0.8839	0.8833	13.4778	13.4081	6.4824	6.4821
Q3	3.6359	3.6340	18.0413	17.9760	13.7234	13.7233
Q4	12.6631	12.6591	23.8262	23.8071	24.9657	24.9658
Q5	82.7012	82.7077	36.3043	36.5071	53.0816	53.0820
90–95	19.2587	19.2582	9.1490	9.1985	13.3950	13.3951
95–99	29.2337	29.2371	8.9201	8.9964	14.5243	14.5244
T1	15.7834	15.7900	2.9787	3.0164	5.4092	5.4093
Wealth Gini	0.7827	0.7828	0.2811	0.2835	0.5149	0.5149
K/Y	11.0678	11.0672	10.3031	10.2993	11.4788	11.4781
ΔC	-2.2794	-2.3179	-1.7865	-1.8358	-1.7581	-1.8120

Figure 6: Impulse responses, KMP calibration.



5.4 Accuracy and efficiency

Table 11 reports the forecasting errors for the dynamic and static forecasts as in Section 4.3. We can see that the forecasting rules in the Xpa algorithm are very accurate and comparable with the KS algorithm. For example, the max DH error is 0.345 in the Xpa algorithm and 0.219 in the KS algorithm for the KMP calibration. We see that the mean DH error tends to be higher in good times (i.e., $z_t = z_h$) in both the Xpa and KS algorithms.

Finally, we display the computation time in Table 12. The Xpa algorithm is a factor of 100–200 times faster than the KS algorithm. While the number of outer loops tends to be higher in the Xpa algorithm, it takes less than 1/100th of a second per loop to calculate.

In our computation, we solve the KMP model within seconds with the Xpa algorithm using Fortran and OpenMP. Note that most of the computation time is in the inner loop.[23]

[23] We can solve the inner loop even faster by using the endogenous grid point method (Carroll, 2006; Barillas and Fernández-Villaverde, 2007; Fella, 2014), for example.

Table 11: DH statistics, RMSE, and R–squared.

		DHmax, %	DHmean, %	RMSE, %	R–squared
KMP					
KS	z_h	0.2069	0.0461	0.0033	1.0000
	z_l	0.2193	0.0062	0.0043	1.0000
Xpa	z_h	0.3404	0.0988	0.0050	1.0000
	z_l	0.3445	0.0168	0.0118	1.0000
KS					
KS	z_h	0.0465	0.0128	0.0010	1.0000
	z_l	0.0502	0.0021	0.0017	1.0000
Xpa	z_h	0.1659	0.1051	0.0047	1.0000
	z_l	0.1611	0.0101	0.0059	1.0000
$+\sigma(y)$					
KS	z_h	0.0385	0.0067	0.0007	1.0000
	z_l	0.0371	0.0016	0.0016	1.0000
Xpa	z_h	0.2994	0.1142	0.0073	1.0000
	z_l	0.2625	0.0079	0.0126	1.0000

Notes: Taking the dynamic and static forecast errors (sorted for each z_i) $\{u_{K,t}\}_{t=1}^{T_i}$ and $\{e_{K,t}\}_{t=1}^{T_i}$ as given, DHmax is the maximum value of $\{u_{K,t}\}_{t=1}^{T_i}$, DHmean is the average of $\{u_{K,t}\}_{t=1}^{T_i}$, RMSE is $\sqrt{\sum_{t=1}^{T_i} e_{K,t}^2}$, and R–squared is $1 - \sum_{t=1}^{T_i} e_{K,t}^2 / \sum_{t=1}^{T_i}(K_{t+1} - \bar{K})^2$ where $\bar{K} = (1/T)\sum_{t=1}^{T_i} K_{t+1}$.

Table 12: Computational time.

	Inner	Outer	# of loop	Total
KMP				
KS	0.455	68.228	14.000	961.568
Xpa	0.360	0.003	14.000	5.078
Xpa/KS	0.792	<0.001	1.000	0.005
KS				
KS	0.203	30.939	10.000	311.421
Xpa	0.301	0.000	15.000	4.520
Xpa/KS	1.484	<0.001	1.500	0.015
$+\sigma(y)$				
KS	0.243	37.082	12.000	447.897
Xpa	0.365	0.001	13.000	4.754
Xpa/KS	1.503	<0.001	1.083	0.011

Notes: Computation using a workstation with Xeon E5–2696v4 (2.2 Ghz, 44 cores).

6 Concluding remarks

In this paper, the Xpa algorithm originally proposed by den Haan and Rendahl (2010) is applied to the heterogeneous–firm models in Khan and Thomas (2003, 2008) and the heterogeneous–household models in Krueger et al. (2016). We demonstrate that the Xpa algorithm is much faster than the KS algorithm and has a similar accuracy in terms of errors in the forecasting rules. Future research may include applying the Xpa algorithm to other classes of heterogeneous–agent models, such as HANK models (e.g., Bayer et al., 2018; Gornemann et al., 2016; Kaplan et al., 2018). As we can solve the models using a projection method that preserves nonlinearity at the aggregate level, structural estimations by matching the micro– and macro–level moments in these models by using the Xpa algorithm may also be of great interest.[24]

[24]See, e.g., Mongey and Williams (2017); Winberry (2018); Williams (2017) for recent attempts at structural estimations using Reiter's (2009) perturbation methods. Note that none of these studies can investigate the higher–order effects of aggregate uncertainty as certainty equivalence holds.

References

Ahn, S., G. Kaplan, B. Moll, , T. Winberry, and C. Wolf (2018): "When Inequality Matters for Macro and Macro Matters for Inequality," *NBER Macroeconomics Annual*, 1–75.

Algan, Y., O. Allais, and W. J. den Haan (2008): "Solving heterogeneous-agent models with parameterized cross-sectional distributions," *Journal of Economic Dynamics and Control*, 32, 875–908.

——— (2010): "Solving the incomplete markets model with aggregate uncertainty using parameterized cross-sectional distributions," *Journal of Economic Dynamics and Control*, 34, 59–68.

Barillas, F. and J. Fernández-Villaverde (2007): "A generalization of the endogenous grid method," *Journal of Economic Dynamics and Control*, 31, 2698–2712.

Bayer, C. and R. Luetticke (2018): "Solving heterogeneous agent models in discrete time with many idiosyncratic states by perturbation methods," Working Paper.

Bayer, C., R. Luetticke, L. Pham-Dao, and V. Tjaden (2018): "Precautionary Savings, Illiquid Assets, and the Aggregate Consequences of Shocks to Household Income Risk," *Econometrica*, forthcoming.

Carroll, C. (2006): "The method of endogenous gridpoints for solving dynamic stochastic optimization problems," *Economics Letters*, 91, 312–320.

Chang, Y. and S.-B. Kim (2007): "Heterogeneity and Aggregation: Implications for Labor-Market Fluctuations," *American Economic Review*, 97, 1939–1956.

——— (2014): "Heterogeneity and Aggregation: Implications for Labor-Market Fluctuations: Reply," *American Economic Review*, 104, 1461–1466.

Childers, D. (2018): "Solution of Rational Expectations Models with Function Valued States," Working paper.

Coeurdacier, N., H. Rey, and P. Winant (2011): "The Risky Steady State," *American Economic Review*, 101, 398–401.

Cooper, R. W. and J. C. Haltiwanger (2006): "On the Nature of Capital Adjustment Costs," *Review of Economic Studies*, 73, 611–633.

den Haan, W. J. (2010): "Comparison of Solutions to the Incomplete Markets Model with Aggregate Uncertainty," *Journal of Economic Dynamics and Control*, 34, 4–27.

den Haan, W. J., K. L. Judd, and M. Juillard (2010): "Computational suite of models with heterogeneous agents: Incomplete markets and aggregate uncertainty," *Journal of Economic Dynamics and Control*, 34, 1–3.

den Haan, W. J. and P. Rendahl (2010): "Solving the Incomplete Markets Model with Aggregate Uncertainty Using Explicit Aggregation," *Journal of Economic Dynamics and Control*, 34, 69–78.

Fella, G. (2014): "A generalized endogenous grid method for non-smooth and non-concave problems," *Review of Economic Dynamics*, 17, 329–344.

Gordon, G. (2011): "Computing Dynamic Heterogeneous-Agent Economies: Tracking the Distribution," PIER Working Paper 11-018.

Gornemann, N. M., K. Kuester, and M. Nakajima (2016): "Doves for the Rich, Hawks for the Poor? Distributional Consequences of Monetary Policy," Board of Governors of the Federal Reserve System International Finance Discussion Papers 1167.

Jang, Y., T. Sunakawa, and M. Yum (2018): "Is Household Heterogeneity Important for Business Cycles?" Working Paper.

Kaplan, G., B. Moll, and G. L. Violante (2018): "Monetary Policy According to HANK," *American Economic Review*, 108, 697–743.

Khan, A. (2017): "Large Recessions in an Overlapping Generations with Unemployment," 2017 Meeting Papers 1559, Society for Economic Dynamics.

Khan, A. and J. K. Thomas (2003): "Nonconvex Factor Adjustments in Equilibrium Business Cycle Models: Do Nonlinearities Matter?" *Journal of Monetary Economics*, 50, 331–360.

——— (2008): "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics," *Econometrica*, 76, 395–436.

Kim, H. (2018): "Inequality, Portfolio Choice, and the Business Cycle," Working Paper.

Krueger, D., K. Mitman, and F. Perri (2015): "Macroeconomics and Household Heterogeneity," Working paper.

——— (2016): "Macroeconomics and Household Heterogeneity," in *Handbook of Macroeconomics*, ed. by J. B. Taylor and H. Uhlig, Elsevier, vol. 2, 843–921.

Krusell, P. and A. A. J. Smith (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106, 867–896.

Maliar, L., S. Maliar, and F. Valli (2010): "Solving the Incomplete Markets Model with Aggregate Uncertainty Using the Krusell-Smith Algorithm," *Journal of Economic Dynamics and Control*, 34, 42–49.

Mongey, S. and J. Williams (2017): "Firm dispersion and business cycles: Estimating aggregate shocks using panel data," Mimeo.

Okahata, N. (2018): "An Alternative Solution Method for Continuous-Time Heterogeneous Agent Models with Aggregate Shocks," Mimeo.

Prescott, E. C. (1986): "Theory Ahead of Business Cycle Measurement," *Federal Reserve Bank of Minneapolis Quarterly Review*, 10, 9–22.

Reiter, M. (2009): "Solving heterogeneous-agent models by projection and perturbation," *Journal of Economic Dynamics and Control*, 33, 649–665.

——— (2010a): "Approximate and Almost-Exact Aggregation in Dynamic Stochastic Heterogeneous-Agent Models," Economics Series 258, Institute for Advanced Studies.

——— (2010b): "Solving the incomplete markets model with aggregate uncertainty by backward induction," *Journal of Economic Dynamics and Control*, 34, 28–35.

Rios-Rull, V. (1999): "Computation of Equilibria in Heterogenous Agent Models," in *Computational Methods for the Study of Dynamic Economies*, Oxford University Press.

Rouwenhorst, K. G. (1995): "Asset Pricing Implications of Equilibrium Business Cycle Models," in *Frontiers of Business Cycle Research*, ed. by T. Cooley, New Jersey: Princeton University Press, 294–330.

Sunakawa, T. (2012): "Applying the Explicit Aggregation Algorithm to Discrete Choice Economies: With an Application to Estimating the Aggregate Technology Shock Process," Mimeo.

Takahashi, S. (2014): "Heterogeneity and Aggregation: Implications for Labor-Market Fluctuations: Comment," *American Economic Review*, 104, 1446–1460.

Tauchen, G. (1986): "Finite state markov-chain approximations to univariate and vector autoregressions," *Economics Letters*, 20, 177–181.

Terry, S. J. (2017): "Alternative Methods for Solving Heterogeneous Firm Models," *Journal of Money, Credit and Banking*, 49, 1081–1111.

Thomas, J. K. (2002): "Is Lumpy Investment Relevant for the Business Cycle?" *Journal of Political Economy*, 110, 508–534.

Williams, J. (2017): "Bayesian estimation of DSGE models with heterogeneous agents," Mimeo.

Winberry, T. (2018): "A Method for Solving and Estimating Heterogeneous Agent Macro Models," *Quantitative Economics*, forthcoming.

Young, E. R. (2005): "Approximate Aggregation," Mimeo, University of Virginia.

——— (2010): "Solving the incomplete markets model with aggregate uncertainty using the Krusell-Smith algorithm and non-stochastic simulations," *Journal of Economic Dynamics and Control*, 34, 36–41.

Appendix (not for publication)

A Details of heterogeneous–firm models

A.1 Stationary equilibrium

We use steady–state information for both the bias correction and the ϵ–indexed aggregation in the Xpa algorithm. As a precise solution of the stationary equilibrium is necessary, we explain how to compute the stationary equilibrium in the following steps.

1. Make an initial guess of w and $p = \eta/w$ because of the household's utility with indivisible labor.

2. Taking p and w as given, solve for the value function and the individual decision rules

   (a) Make an initial guess of the value function $v^{(0)}(\epsilon_i, k)$.[25]

   (b) Taking $v^{(l-1)}(\epsilon_i, k)$ as given (l is an index for iteration in the inner loop), approximate the expected value function $\sum_{\epsilon'|\epsilon_i} \pi_\epsilon(\epsilon'|\epsilon_i) v^{(l)}(\epsilon', k')$ and solve

$$E_0(\epsilon_i) = \max_{k'>0} \left\{ -\gamma p k' + \beta \sum_{\epsilon'|\epsilon_i} \pi_\epsilon(\epsilon'|\epsilon_i) v^{(l)}(\epsilon', k') \right\}, \tag{14}$$

$$E_1(\epsilon_i, k) = \max_{k' \in \Omega(k)} \left\{ -\gamma p k' + \beta \sum_{\epsilon'|\epsilon_i} \pi_\epsilon(\epsilon'|\epsilon_i) v^{(l)}(\epsilon', k') \right\}, \tag{15}$$

   for $k^*(\epsilon_i)$ and $k^c(\epsilon_i, k)$. The individual decision rules are given by

$$\bar{g}_k(\epsilon_i, k) = \alpha(\epsilon_i, k) k^*(\epsilon_i) + (1 - \alpha(\epsilon_i, k)) k^c(\epsilon_i, k),$$
$$\bar{g}_y(\epsilon_i, k) = \epsilon_i F(k, n^*(\epsilon_i, k)),$$

   where $n^*(\epsilon_i, k) = \arg\max_n \{\epsilon_i F(k, n) - wn\}$.

[25] k is a continuous variable and we discretize it on a computer by grid points. We suppose that k takes a value of grid points $k_j \in \{k_1, ..., k_{n_k}\}$ and use unidimensional piecewise cubic splines to interpolate the values between the grid points. We omit the index for k for the sake of exposition.

(c) Update the value function $v^{(l)}(\epsilon_i, k)$ by

$$\begin{aligned} v^{(l)}(\epsilon_i, k) &= p \max_n \{\epsilon_i F(k, n) - wn\} + p(1-\delta)k \\ &- \eta \int_0^{\hat{\xi}(\epsilon_i, k)} \xi G'(\xi) d\xi \\ &+ \alpha(\epsilon_i, k) E_0(\epsilon_i) + (1 - \alpha(\epsilon_i, k)) E_1(\epsilon_i, k), \end{aligned} \tag{16}$$

where

$$\begin{aligned} \hat{\xi}(\epsilon_i, k) &= (E_0(\epsilon_i) - E_1(\epsilon_i, k)) / \eta, \\ \alpha(\epsilon_i, k) &= G\left(\hat{\xi}(\epsilon_i, k)\right). \end{aligned}$$

(d) Iterate Steps (b) and (c) until $\left\| v^{(l)} - v^{(l-1)} \right\|$ is sufficiently small.

3. Solve for the stationary distribution $\bar{\mu}(\epsilon_i, k)$ implied by the individual decision rules (see below).

4. Update the prices: Given the stationary distribution, the individual decision rules are aggregated as

$$Y = \sum_i \int \bar{g}_y(\epsilon_i, k) \bar{\mu}(\epsilon_i, k) dk, \tag{17}$$

$$K = \sum_i \int \bar{g}_k(\epsilon_i, k) \bar{\mu}(\epsilon_i, k) dk, \tag{18}$$

$C = Y - \delta K$ and $p = C^{-1}$.

5. Iterate Steps 2–4 until convergence.

In Step 3, we approximate the dynamics of $\bar{\mu}(\epsilon_i, k)$ using the non–stochastic method with histogram (Young, 2010). That is, $\bar{\mu}(\epsilon_i, k)$ is a continuous object with regard to k and we suppose that k takes a value of grid points $k_j \in \{k_1, ..., k_{\tilde{n}_k}\}$.[26] Given the current period's distribution $\mu(\epsilon_i, k_j)$

[26] We have finer grid points for approximating the distribution than interpolating the value function in Step 2, i.e., $\tilde{n}_k > n_k$.

for $i = 1, ..., n_e$ and $j = 1, ..., \tilde{n}_k$, the next period's distribution $\mu(\epsilon_{i'}, k_{j'})$ for $i' = 1, ..., n_e$ and $j' = 1, ..., \tilde{n}_k$ is obtained from the transition probabilities from the current period's index (i, j) to the next period's index (i', j'). There are two cases of transition depending on the value of a fixed cost drawn by the firms indexed by (i, j). (i) When a firm pays the fixed cost $\xi \leq \hat{\xi}(\epsilon_i, k_j)$ with probability $\alpha(\epsilon_i, k_j)$: for each i, there is an index j_0' such that $k' = k^*(\epsilon_i) \in [k_{j_0'}, k_{j_0'+1}]$.[27] Then the transition probabilities from (i, j) to (i', j_0') and $(i', j_0' + 1)$ are

$$\begin{aligned}\mathcal{P}(i', j_0'|i, j) =& \pi_\epsilon(\epsilon_{i'}|\epsilon_i)\omega^*(\epsilon_i)\alpha(\epsilon_i, k_j), \\ \mathcal{P}(i', j_0' + 1|i, j) =& \pi_\epsilon(\epsilon_{i'}|\epsilon_i)\left(1 - \omega^*(\epsilon_i)\right)\alpha(\epsilon_i, k_j),\end{aligned}$$

where $\pi_\epsilon(\epsilon_{i'}|\epsilon_i)$ is the transition probability from i to i' and $\omega^*(\epsilon_i) = \frac{k_{j_0'+1} - k^*(\epsilon_i)}{k_{j_0'+1} - k_{j_0'}}$. (ii) When a firm does not pay the fixed cost $\xi > \hat{\xi}(\epsilon_i, k_j)$ with probability $1 - \alpha(\epsilon_i, k)$: for each (i, j), there is an index j_1' such that $k' = k^c(\epsilon_i, k_j) \in [k_{j_1'}, k_{j_1'+1}]$. Then the transition probabilities from (i, j) to (i', j_1') and $(i', j_1' + 1)$ are

$$\begin{aligned}\mathcal{P}(i', j_1'|i, j) =& \pi_\epsilon(\epsilon_{i'}|\epsilon_i)\omega^c(\epsilon_i, k_j)\left(1 - \alpha(\epsilon_i, k_j)\right), \\ \mathcal{P}(i', j_1' + 1|i, j) =& \pi_\epsilon(\epsilon_{i'}|\epsilon_i)\left(1 - \omega^c(\epsilon_i, k_j)\right)\left(1 - \alpha(\epsilon_i, k_j)\right),\end{aligned}$$

where $\omega^c(\epsilon_i, k_j) = \frac{k_{j_1'+1} - k^c(\epsilon_i, k_j)}{k_{j_1'+1} - k_{j_1'}}$. Given the transition probabilities for all (i, j), we have a transition matrix $\mathcal{P}$ such that $\vec{\mu} = \mathcal{P}\vec{\mu}$ holds, where $\vec{\mu}$ is a stacked $(n_e\tilde{n}_k \times 1)$ vector of the stationary distribution. $\vec{\mu}$ is obtained by an iterative method with an initial distribution or eigenvalue and eigenvector decomposition of the matrix $\mathcal{P}$.[28]

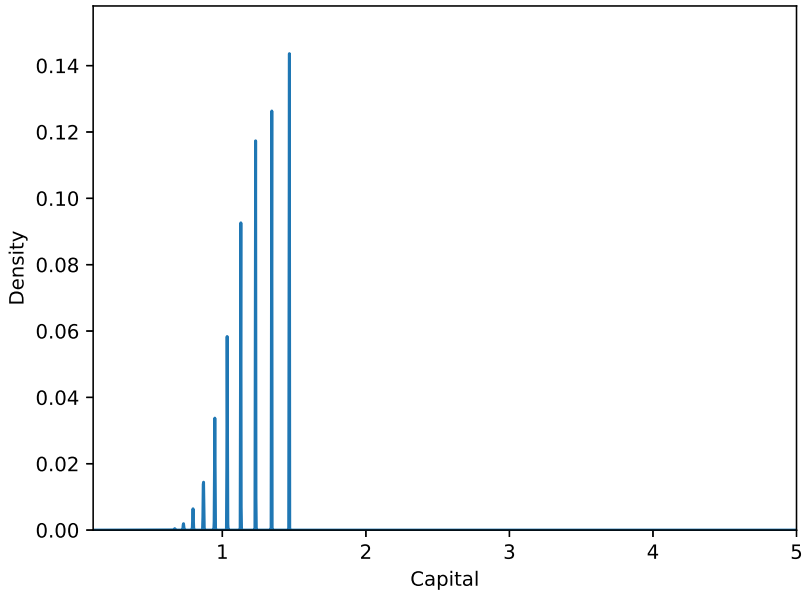
As shown in Figure 7, the distribution takes the form of a discrete function, especially in the traditional model of KT (2003) without idiosyncratic productivity shocks and constrained investment. Therefore, we need fine grid points to approximate such a function.

[27] Note that all the firms indexed by i choose the same level of capital regardless of k_j.

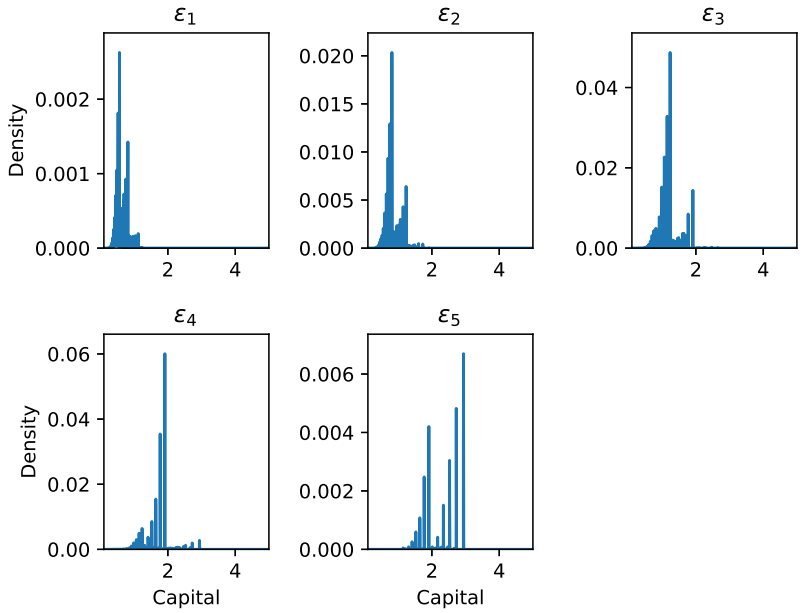
[28] $\mathcal{P}$ is a sparse matrix. To compute the eigenvalue and eigenvector decomposition, Reiter (2009) suggests using MATLAB's command eigs. We implement its functionality in Fortran by using the libraries arpack and Sparse BLAS.

Figure 7: Stationary distributions.

a. Traditional model



b. Extended model



A.2 Xpa and KS algorithms

In both the Xpa and KS algorithms, we solve for the individual decision rule g and the aggregate forecasting rule Γ by the following nested loops. In the *inner* loop, given the forecasting rule $\Gamma^{(n-1)}$ (n is an index for iteration), we solve for the individual decision rule $g^{(n)}$. In the *outer* loop, given the individual decision rule obtained in the inner loop, we solve for the forecasting rule $\Gamma^{(n)}$. The inner and outer loops are iterated until both g and Γ converge. At the convergence, the individual decision rule is consistent with the forecasting rule. Each KS and Xpa algorithm achieves consistency in a different manner. The inner loop is common to both algorithms and only the outer loop differs.

Inner loop Taking as given the forecasting rules $K_{t+1} = \Gamma_K^{(n-1)}(z_t, K_t)$ and $p_t = \Gamma_p^{(n-1)}(z_t, K_t)$, the common inner loop for the two algorithms is as follows.

1. Make an initial guess of the value function $v^{(0)}(\epsilon_i, k; z_t, K_t)$.[29]

2. Taking $v^{(l-1)}(\epsilon_i, k; z_t, K_t)$ and $K_{t+1} = \Gamma_K^{(n-1)}(z_t, K_t)$ as given (l is an index for iteration in the inner loop),

   (a) Approximate the expected value function in the next period

   $$h_{\epsilon_i z_t}^{(l-1)}(k', K_t) \simeq \sum_{\epsilon'} \pi_\epsilon(\epsilon'|\epsilon_i) \sum_{z_{t+1}} \pi_z(z_{t+1}|z_t) v^{(l-1)}(\epsilon', k'; z_{t+1}, \Gamma_K^{(n-1)}(z_t, K_t))$$

   for each ϵ_i and z_t.

   (b) Solve

   $$E_0(\epsilon_i; z_t, K_t) = \max_{k'>0} \left\{ -\gamma p_t k' + \beta h_{\epsilon_i z_t}^{(l-1)}(k', K_t) \right\}, \tag{19}$$

   $$E_1(\epsilon_i, k; z_t, K_t) = \max_{k' \in \Omega(k)} \left\{ -\gamma p_t k' + \beta h_{\epsilon_i z_t}^{(l-1)}(k', K_t) \right\}, \tag{20}$$

[29] (k, K_t) are continuous variables and we discretize them on a computer by a rectangle of grid points. We suppose that k takes a value of $k_j \in \{k_1, ..., k_{n_k}\}$ and K_t takes a value of $K_m \in \{K_1, ..., K_{n_K}\}$ and use bidimensional piecewise cubic splines to interpolate the values between the grid points.

for $k^*(\epsilon_i; z_t, K_t)$ and $k^c(\epsilon_i, k; z_t, K_t)$. The individual decision rules are given by

$$\begin{aligned} g_k(\epsilon_i, k; z_t, K_t) &= \alpha(\epsilon_i, k; z_t, \mu_t) k^*(\epsilon_i; z_t, K_t) \\ &\quad + (1 - \alpha(\epsilon_i, k; z_t, K_t)) k^c(\epsilon_i, k; z_t, K_t), \\ g_y(\epsilon_i, k; z_t, K_t) &= \epsilon_i z_t F(k, n^*(\epsilon_i, k, z_t, K_t)), \end{aligned}$$

where $n^*(\epsilon_i, k, z_t, K_t) = \arg\max_n \{\epsilon_i z_t F(k, n) - w_t n\}$.

3. Update the value function $v^{(l)}(\epsilon_i, k; z_t, K_t)$ by

$$\begin{aligned} v^{(l)}(\epsilon_i, k; z_t, K_t) &= p_t \max_n \{\epsilon_i z F(k, n) - w_t n\} + p_t(1 - \delta)k \\ &- \eta \int_0^{\hat{\xi}(\epsilon_i, k; z_t, K_t)} \xi G'(\xi) d\xi + \alpha(\epsilon_i, k; z_t, K_t) E_0(\epsilon_i; z_t, K_t) \\ &+ (1 - \alpha(\epsilon_i, k; z_t, K_t)) E_1(\epsilon_i, k; z_t, K_t), \end{aligned}$$

where

$$\begin{aligned} \hat{\xi}(\epsilon_i, k; z_t, K_t) &= (E_0(\epsilon_i; z_t, K_t) - E_1(\epsilon_i, k; z_t, K_t)) / (w_t p_t), \\ \alpha(\epsilon_i, k; z_t, K_t) &= G\left(\hat{\xi}(\epsilon_i, k; z_t, K_t)\right). \end{aligned}$$

4. Iterate Steps 2 and 3 until $\left\| v^{(l)} - v^{(l-1)} \right\|$ is sufficiently small.

Outer loop in the KS algorithm In the outer loop in the KS algorithm, the individual decision rules aggregate with $\mu_t(\epsilon_i, k_j)$ in each period. The model is simulated for long enough to obtain the sequence of aggregate variables.

Taking the value function $v(\epsilon_i, k; z_t, K_t)$ obtained in the inner loop as given, the outer loop in the KS algorithm is as follows.

1. Make a sequence of $\{z_t\}$ for $t = 1, ..., T$ and an initial distribution $\mu_1(\epsilon_i, k_j)$ approximated by a histogram at each grid point for $i = 1, ..., n_e$ and $j = 1, ..., \tilde{n}_k$ so that $\sum_i \sum_j \mu_1(\epsilon_i, k_j) = 1$ holds.

2. In each period $t = 1, ..., T$, z_t and the current period's distribution $\mu_t(\epsilon_i, k_j)$ are given.

   (a) The aggregate capital is given by $K_t = \sum_i \sum_j k_j \mu_t(\epsilon_i, k_j)$. Taking $v(\epsilon_i, k_j; z_t, K_t)$ and $K_{t+1} = \Gamma_K^{(n-1)}(z_t, K_t)$ as given, make a conjecture of $\tilde{p}_t$ (which is not necessarily the forecasted price) and solve for $g_y(\epsilon_i, k_j; z_t, K_t)$ and $g_k(\epsilon_i, k_j; z_t, K_t)$ by the same procedure as in Step 2 for the inner loop. Then aggregate the individual decision rules with the distribution to obtain

$$\begin{aligned} Y_t &= \sum_i \sum_j g_y(\epsilon_i, k_j; z_t, K_t)\mu_t(\epsilon_i, k_j), \\ K_{t+1} &= \sum_i \sum_j g_k(\epsilon_i, k_j; z_t, K_t)\mu_t(\epsilon_i, k_j), \end{aligned}$$

   $C_t = Y_t - \gamma K_{t+1} + (1-\delta)K_t$, and $p_t = C_t^{-1}$. That is, there is a mapping from the conjectured price to the new price and it is denoted by $p_t = F(\tilde{p}_t)$.

   (b) Iterate Step 2(b) to find the market–clearing price p_t^* satisfying $p_t^* = F(p_t^*)$.

   (c) Update the next period's distribution $\mu_{t+1}(\epsilon_i, k_j)$ by $\vec{\mu}_{t+1} = \mathcal{P}\vec{\mu}_t$ where $\vec{\mu}_t$ is a stacked vector of the distribution in period t and $\mathcal{P}$ is a transition matrix with the transition probabilities from the current period's index (i, j) to the next period's index (i', j') and $(i', j'+1)$. There are two cases. When $\xi \leq \hat{\xi}(\epsilon_i, k; z_t, K_t)$,

$$\begin{aligned} \mathcal{P}(i', j_0'|i, j) =& \pi_\epsilon(\epsilon_{i'}|\epsilon_i)\omega^*(\epsilon_i; z_t, K_t)\alpha(\epsilon_i, k_j; z_t, K_t), \\ \mathcal{P}(i', j_0'+1|i, j) =& \pi_\epsilon(\epsilon_{i'}|\epsilon_i)\left(1 - \omega^*(\epsilon_i; z_t, K_t)\right)\alpha(\epsilon_i, k_j; z_t, K_t), \end{aligned}$$

   where $\omega^*(\epsilon_i; z_t, K_t) = \frac{k_{j_0'+1} - k^*(\epsilon_i; z_t, K_t)}{k_{j_0'+1} - k_{j_0'}}$ and when $\xi > \hat{\xi}(\epsilon_i, k; z_t, K_t)$,

$$\begin{aligned} \mathcal{P}(i', j_1'|i, j) =& \pi_\epsilon(\epsilon_{i'}|\epsilon_i)\omega^c(\epsilon_i, k_j; z_t, K_t)\left(1 - \alpha(\epsilon_i, k_j; z_t, K_t)\right), \\ \mathcal{P}(i', j_1'+1|i, j) =& \pi_\epsilon(\epsilon_{i'}|\epsilon_i)\left(1 - \omega^c(\epsilon_i, k_j; z_t, K_t)\right)\left(1 - \alpha(\epsilon_i, k_j; z_t, K_t)\right), \end{aligned}$$

   where $\omega^c(\epsilon_i, k_j; z_t, K_t) = \frac{k_{j_1'+1} - k^c(\epsilon_i, k_j; z_t, K_t)}{k_{j_1'+1} - k_{j_1'}}$.

3. Update the forecasting rules $K_{t+1} = \Gamma_K^{(n)}(z_t, K_t)$ and $p_t = \Gamma_p^{(n)}(z_t, K_t)$ by using the sequence of $\{z_t, K_t, p_t\}_{t=t_0+1}^{T}$ obtained in Step 2 for $t = 1, ..., T$ (the first t_0 periods are discarded). Specifically, using the sequence obtained, we construct a subset of the sequence sorted by the realization of z_t for each $z_i \in \{z_1, ..., z_{n_z}\}$ and estimate the forecasting rule by ordinary least squares in the form of

$$\log K_{t+1} = b_{K,0}(z_i) + b_{K,1}(z_i) \log K_t,$$
$$\log p_t = b_{p,0}(z_i) + b_{p,1}(z_i) \log K_t,$$

for each $i = 1, ..., n_z$.

Outer loop in the Xpa algorithm In the Xpa algorithm, aggregate technology z_m and aggregate capital K_m are given as a grid point for $m = 1, ..., n_e n_K$.[30] Taking the expected value functions as given, the fixed point problem $p_m = F(p_m)$ is solved as in the KS algorithm but at each grid point. The market clearing price is solved for $n_z n_K$ grid points in the Xpa algorithm, whereas it is solved in each period of simulations for T periods in the KS algorithm. The Xpa algorithm is faster than the KS algorithm as $n_z n_K << T$ holds.

Taking the value function $v(\epsilon_i, k; z, K)$ obtained in the inner loop as given, the outer loop in the Xpa algorithm is as follows.

1. Calculate the bias–correction terms

$$\zeta_k(\epsilon_i) = K_{ss}(\epsilon_i) - \bar{g}_k(K_{ss}(\epsilon_i)),$$
$$\zeta_y(\epsilon_i) = Y_{ss}(\epsilon_i) - \bar{g}_y(K_{ss}(\epsilon_i)),$$

where $K_{ss}(\epsilon_i) = \sum_j k_j \bar{\mu}_k(k_j|\epsilon_i) = \sum_j k_j \bar{\mu}(\epsilon_i, k_j)/\phi(\epsilon_i)$. Also compute $\psi(\epsilon_i) = K_{ss}(\epsilon_i)/K_{ss}$ for the ϵ–indexed aggregation.

[30] We explicitly consider the aggregate state variables (z_m, K_m) as grid points and use subscripts m instead of t. The next period's aggregate capital is denoted by K' instead of K_{t+1}. The other aggregate variables are also computed at each grid point, so they are indexed by m rather than time t.

2. At each grid point (z_m, K_m) indexed by $m = 1, ..., n_z n_K$,

   (a) The aggregate capital K_m is given as a grid point.

   (b) Taking $v(\epsilon_i, k; z_m, K_m)$ and $K'_m = \Gamma_K^{(n-1)}(z_m, K_m)$ as given, make a conjecture of $\tilde{p}_m$ (which is not necessarily the forecasted price) and solve for $g_y(\epsilon_i, k; z_m, K_m)$ and $g_k(\epsilon_i, k; z_m, K_m)$ by the same procedure as in Step 2 for the inner loop. Then evaluate the individual decision rules at $k = K(\epsilon_i) \approx \psi(\epsilon_i)K_m$ for each ϵ_i and calculate the weighted average to obtain

$$\begin{aligned} Y_m &= \sum_i \left[g_y(\epsilon_i, \psi(\epsilon_i)K_m; z_m, K_m) + \zeta_y(\epsilon_i) \right] \phi(\epsilon_i), \\ K'_m &= \sum_i \left[g_k(\epsilon_i, \psi(\epsilon_i)K_m; z_m, K_m) + \zeta_k(\epsilon_i) \right] \phi(\epsilon_i), \end{aligned}$$

   $C_m = Y_m - \gamma K'_m + (1-\delta)K_m$, and $p_m = C_m^{-1}$. That is, there is a mapping denoted by $p_m = F(\tilde{p}_m)$.

   (c) Iterate Step 2(b) to find the market–clearing price p_m^* satisfying $p_m^* = F(p_m^*)$.

3. Update the forecasting rules $K' = \Gamma_K^{(n)}(z, K)$ and $p = \Gamma_p^{(n)}(z, K)$ by using (K'_m, p_m) obtained in Step 2 at each grid point (z_m, K_m).

Numerical setup The individual capital k ranges between [0.1,5.0] and is spaced into $n_k = 101$ grid points. The aggregate capital K ranges between $[0.75K_{ss}, 1.25K_{ss}]$ around the steady state value of K_{ss} and is evenly spaced into $n_K = 5$ grid points. The distribution of k is approximated by $\tilde{n}_k = 2001$ grid points by the non–stochastic method with histogram (Young, 2010). For the idiosyncratic and aggregate shocks, $n_e = 5$ and $n_z = 5$ grid points are used with Tauchen's (1986) method to approximate the AR(1) processes. We set the grid points to cover 99% (95%) of the distribution implied by the original AR(1) process for the idiosyncratic (aggregate) productivity shock.[31] A two–dimensional cubic spline is used to approximate the expected value function. To solve for k^* and k^c in (19) and (20), a variant of Newton's method or the golden section search

[31] That is, $\epsilon_{n_e} = -\epsilon_1 = 2.58\sigma_\epsilon/\sqrt{1-\rho_\epsilon^2}$ and $z_{n_z} = -z_1 = 1.96\sigma_z/\sqrt{1-\rho_z^2}$.

method is used. Convergence criteria are 10^{-4} for the outer loop, 10^{-5} for the inner loop, 10^{-6} for market clearing in the outer loop, and 10^{-10} for the other optimizations. The initial forecasting rules are from the social planner's problem. In the KS algorithm, the model is simulated for 2,500 periods and the first 500 periods are discarded. The same sequence of $\{z_t\}$ throughout all the outer loops is used to avoid chatter, which prevents the algorithm from converging within a finite sample.

A.3 More numerical results

Table 13: Business cycle statistics, traditional model.

	Y	C	I	N	K	Z
Standard deviation						
KS	(2.5293)	0.4329	3.6109	0.6680	0.5169	0.5981
Xpa	(2.5134)	0.4370	3.5901	0.6619	0.5142	0.6019
Output correlation						
KS	1.0000	0.8762	0.9740	0.9379	0.0451	0.9995
Xpa	1.0000	0.8807	0.9740	0.9374	0.0447	0.9995

	Persist.	Std.dev.	Skewness	Exc.kur.
Agg. investment rate, I/K				
KS	0.6571	0.0108	0.0747	-0.0968
Xpa	0.6590	0.0106	0.0619	-0.0946

Figure 8: Unconditional business cycle simulation, traditional model.

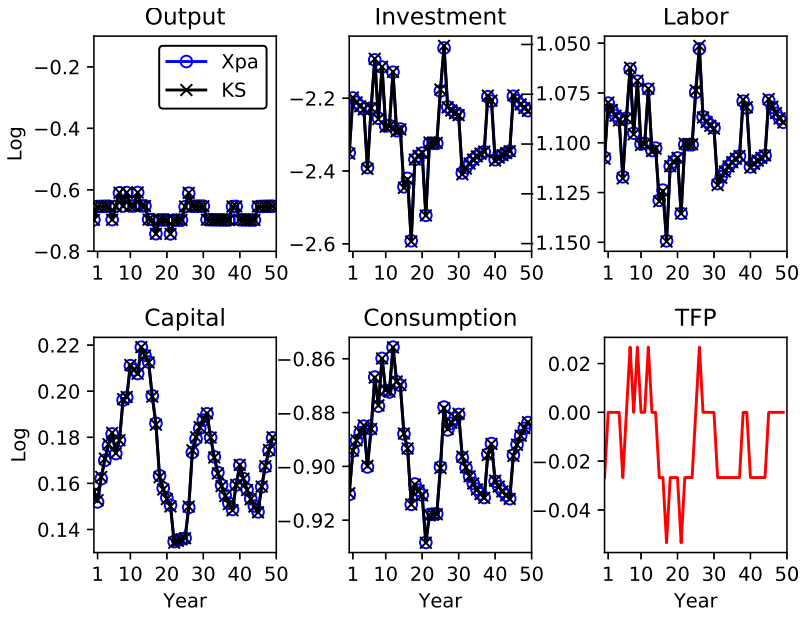


Figure 9: Impulse responses, traditional model.

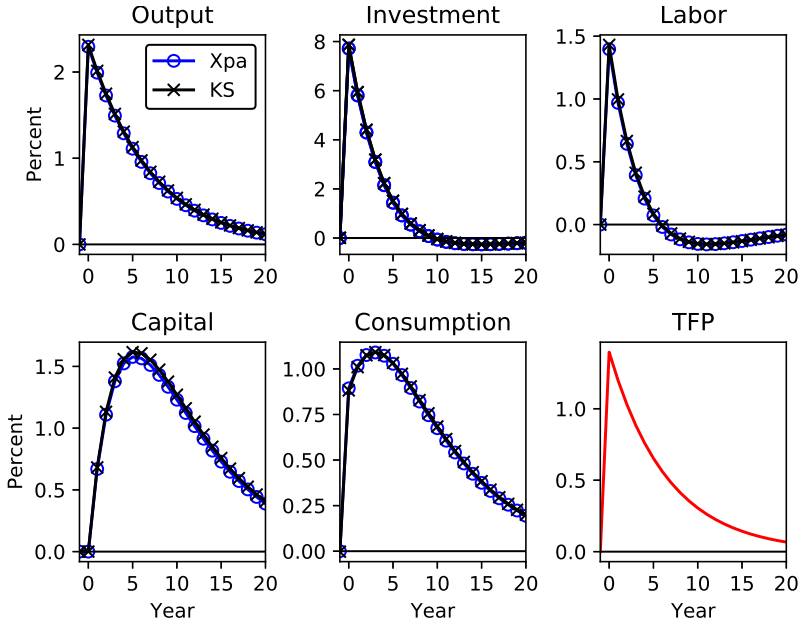


Table 14: Level differences between KS and Xpa, traditional model.

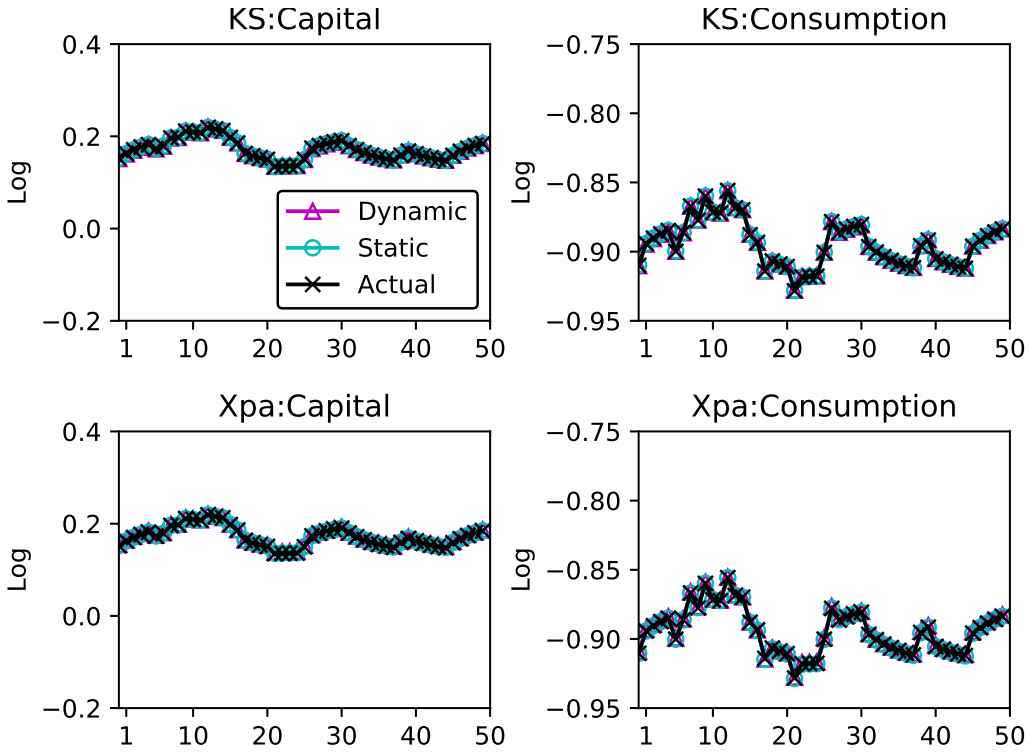
a. Aggregate variables

	K/Y	N	$p = C^{-1}$
KS	2.3476	0.3337	2.4002
Xpa	2.3489	0.3337	2.3998
SS	2.3487	0.3337	2.3998

b. Firm–level investment rates

	Mean	Std.dev.	Inaction	Spike+	Spike-	Invest+	Invest-
KS	0.1046	0.0214	0.2115	0.1279	0.0000	0.7885	0.0000
Xpa	0.1046	0.0214	0.2115	0.1280	0.0000	0.7885	0.0000
SS	0.1046	0.0214	0.2113	0.1284	0.0000	0.7887	0.0000

Figure 10: den Haan fundamental plots, traditional model.



B Details of heterogeneous–household models

B.1 Stationary equilibrium

We explain how to compute the stationary equilibrium in the following steps.

1. Make an initial guess of K and obtain $w = (1-\alpha)(K/L)^{-\alpha}$ and $r = \alpha(K/L)^{\alpha-1} - \delta$ where $L = 1 - (\Pi_l u(z_l) + (1-\Pi_l)u(z_h))$.

2. Taking K, w and r as given, solve for the value function and the individual decision rules

   (a) Make an initial guess of the value function $v^{(0)}(x_i, k)$.

   (b) Taking $v^{(l-1)}(\epsilon_i, k)$ as given (l is an index for iteration in the inner loop), approximate the expected value function and solve

$$\begin{aligned} v^{(l)}(x_i,k) &= \max_{k'\geq 0}\Big\{u\left((1-\tau)wy_ie_i + (1+r)k/\theta - k'\right) \\ &+ \theta\beta_i \sum_{e'}\sum_{y'} \pi_e(e'|e_i)\pi_y(y'|y_i)v^{(l-1)}(e',y',\beta_i,k')\Big\}, \end{aligned}$$

   for the individual decision rule $k' = \bar{g}(x_i, k)$.

   (c) Iterate Step (b) until $\left\|v^{(l)} - v^{(l-1)}\right\|$ is sufficiently small.

3. Solve for the stationary distribution $\bar{\mu}(x_i, k)$ implied by the individual decision rules.

4. Given the stationary distribution, the individual decision rule is aggregated as

$$K = \sum_i \int \bar{g}(x_i,k)\theta\bar{\mu}(x_i,k)dk,$$

   where θ is the survival probability from period t to period $t+1$. Note that the replacing newborns have no assets.

5. Iterate Steps 2–4 until convergence.

The transition matrix for employment status is conditioned on (z_t, z_{t+1}) and denoted by

$$\pi_e(e_{i'}|e_i, z_{t+1}, z_t) = \begin{bmatrix} p_{z_t z_{t+1}} & 1 - p_{z_t z_{t+1}} \\ 1 - q_{z_t z_{t+1}} & q_{z_t z_{t+1}} \end{bmatrix},$$

where $p_{z_t z_{t+1}}$ is the conditional probability to continue to be unemployed and $q_{z_t z_{t+1}}$ is the conditional probability to continue to be employed. Then the transition matrix for the stationary equilibrium is given by

$$\pi_e(e_{i'}|e_i) = \begin{bmatrix} p & 1 - p \\ 1 - q & q \end{bmatrix},$$

where

$$\begin{aligned} p &= \frac{(1 - \Pi_l)u(z_h)}{u_{ss}} \left(\rho_h p_{z_h z_h} + (1 - \rho_h) p_{z_h z_l}\right) + \frac{\Pi_l u(z_l)}{u_{ss}} \left((1 - \rho_l) p_{z_l z_h} + \rho_l p_{z_l z_l}\right), \\ q &= \frac{(1 - \Pi_l)(1 - u(z_h))}{1 - u_{ss}} \left(\rho_h q_{z_h z_h} + (1 - \rho_h) q_{z_h z_l}\right) + \frac{\Pi_l (1 - u(z_l))}{1 - u_{ss}} \left((1 - \rho_l) q_{z_l z_h} + \rho_l q_{z_l z_l}\right), \end{aligned}$$

and $u_{ss} = (1 - \Pi_l)u(z_h) + \Pi_l u(z_l)$. Note that the stationary distribution of employment status implied by $\pi_e(e'|e)$ coincides with the employment rate $1 - u_{ss}$.

In Step 3, given the current period's distribution $\mu(x_i, k_j)$ for $i = 1, ..., n_x$ and $j = 1, ..., \tilde{n}_k$, the next period's distribution $\mu(\epsilon_{i'}, k_{j'})$ for $i' = 1, ..., n_e$ and $j' = 1, ..., \tilde{n}_k$ is obtained from transition probabilities from the current period's index (i, j) to the next period's index (i', j'). With probability $(1 - \theta)$, households die and newly born households with the same properties $x_{i'} = (\epsilon_{i'}, y_{i'}, \beta_i)$ and no assets replace the old. The transition probability from (i, j) to $(i', 1)$ is

$$\mathcal{P}(i', 1|i, j) = \pi_e(e_{i'}|e_i)\pi_y(y_{i'}|y_i)(1 - \theta).$$

With probability θ, households continue to live in the next period and choose $k' = \bar{g}(x_i, k_j)$. Then

the transition probabilities from (i,j) to (i',j') and $(i',j'+1)$ are

$$\begin{aligned}\mathcal{P}(i',j'|i,j) =& \pi_e(e_{i'}|e_i)\pi_y(y_{i'}|y_i)\omega^*(x_i,k_j)\theta,\\ \mathcal{P}(i',j'+1|i,j) =& \pi_e(e_{i'}|e_i)\pi_y(y_{i'}|y_i)\left(1-\omega^*(x_i,k_j)\right)\theta,\end{aligned}$$

where $\omega^*(x_i,k_j)=\frac{k_{j'_0+1}-\bar{g}(x_i,k_j)}{k_{j'_0+1}-k_{j'_0}}$. If $j'=1$, the transition probability from (i,j) to $(i',1)$ is the sum of two probabilities. That is, $\mathcal{P}(i',1|i,j)=\pi_e(e_{i'}|e_i)\pi_y(y_{i'}|y_i)\left(\omega^*(x_i,k_j)\theta+(1-\theta)\right)$.

B.2 Xpa and KS algorithms

Inner loop Taking as given the forecasting rule $K_{t+1}=\Gamma^{(n-1)}(z_t,K_t)$, the common inner loop for the two algorithms is as follows.

1. Make an initial guess of the value function $v^{(0)}(x_i,k;z_t,K_t)$.
2. Taking $v^{(l-1)}(x_i,k;z_t,K_t)$ and $K_{t+1}=\Gamma^{(n-1)}(z_t,K_t)$ as given (l is an index for iteration in the inner loop),
   (a) Approximate the expected value function in the next period

$$\begin{aligned}h^{(l-1)}_{x_iz_t}(k',K_t)\simeq&\sum_{z_{t+1}}\sum_{e'}\sum_{y'}\pi_z(z_{t+1}|z_t)\pi_e(e'|e_i,z_{t+1},z_t)\pi_y(y'|y_i)\\ &\times v^{(l-1)}(e',y',\beta_i,k';z_{t+1},\Gamma(z_t,K_t))\end{aligned}$$

   for each x_i and z_t.

   (b) Solve

$$v^{(l)}(x_i,k;z_t,K_t) = \max_{c,k'\geq 0}\left\{u(c)+\theta\beta_i h^{(l-1)}_{x_iz_t}(k',K_t)\right\}$$

   subject to the budget constraint

$$c+k' = (1-\tau(z_t;\rho))w_ty_ie_i+(1+r_t)k/\theta,$$

where $w_t = (1-\alpha)z_t(K_t/L(z_t))^{-\alpha}$ and $r_t = \alpha(K_t/L(z_t))^{\alpha-1} - \delta$. Note that $L(z_t) = 1 - u(z_t)$ is exogenously given.

3. Iterate Step 2 until $\left\|v^{(l)} - v^{(l-1)}\right\|$ is sufficiently small.

Outer loop in the KS algorithm Taking the value function $v(x_i, k; z_t, K_t)$ obtained in the inner loop as given, the outer loop in the KS algorithm is as follows.

1. Make a sequence of $\{z_t\}$ for $t = 1, ..., T$ and an initial distribution $\mu_1(x_i, k_j)$ approximated by a histogram at each grid point for $i = 1, ..., n_x$ and $j = 1, ..., \tilde{n}_k$ so that $\sum_i \sum_j \mu_1(x_i, k_j) = 1$ holds.

2. In each period $t = 1, ..., T$, z_t and the current period's distribution $\mu_t(x_i, k_j)$ are given.

   (a) The aggregate capital is given by $K_t = \sum_i \sum_j k_j \mu_t(x_i, k_j)$.[32]

   (b) Taking $v(x_i, k_j; z_t, K_t)$ and $K_{t+1} = \Gamma^{(n-1)}(z_t, K_t)$ as given, solve for $g(x_i, k_j; z_t, K_t)$ by the same procedure as in Step 2 for the inner loop. Then aggregate the individual decision rules with the distribution to obtain

$$K_{t+1} = \sum_i \sum_j g(x_i, k_j; z_t, K_t)\theta\mu_t(x_i, k_j).$$

   Update the next period's distribution $\mu_{t+1}(x_i, k_j)$ by $\vec{\mu}_{t+1} = \mathcal{P}\vec{\mu}_t$ where $\vec{\mu}_t$ is a stacked vector of the distribution in period t and $\mathcal{P}$ is a transition matrix with the transition probabilities. With probability $(1-\theta)$, the transition probability from (i, j) to $(i', 1)$ is

$$\mathcal{P}(i', 1|i, j) = \pi_e(e_{i'}|e_i, z_{t+1}, z_t)\pi_y(y_{i'}|y_i)(1-\theta).$$

[32] Note this may not be on the grid points. Therefore we need to approximate $g(\epsilon_i, k_j; z_t, K_t)$ obtained in the inner loop or redo the optimization.

With probability θ, the transition probabilities from (i,j) to (i',j') and $(i',j'+1)$ are

$$\begin{aligned}\mathcal{P}(i',j'|i,j) =& \pi_e(e'|e_i,z_{t+1},z_t)\pi_y(y'|y_i)\omega(x_i,k_j;z_t,K_t)\theta,\\ \mathcal{P}(i',j'+1|i,j) =& \pi_e(e'|e_i,z_{t+1},z_t)\pi_y(y'|y_i)\left(1-\omega(x_i,k_j;z_t,K_t)\right)\theta,\end{aligned}$$

where $\omega(x_i,k_j;z_t,K_t)=\frac{k_{j'+1}-g(x_i,k_j;z_t,K_t)}{k_{j'+1}-k_{j'}}$.

3. Update the forecasting rule $K_{t+1}=\Gamma_K^{(n)}(z_t,K_t)$ by using the sequence of $\{z_t,K_t\}_{t=t_0+1}^{T}$ obtained in Step 2 for $t=1,...,T$ (the first t_0 periods are discarded). Specifically, using the sequence obtained, we construct a subset of the sequence sorted by the realization of z_t for each $z_i\in\{z_1,...,z_{n_z}\}$ and estimate the forecasting rule by ordinary least squares in the form of

$$\log K_{t+1}=b_{K,0}(z_i)+b_{K,1}(z_i)\log K_t,$$

for each $i=1,...,n_z$.

Outer loop in the Xpa algorithm Taking the value function $v(x_i,k;z,K)$ obtained in the inner loop as given, the outer loop in the Xpa algorithm is as follows.

1. Calculate the bias–correction terms

$$\zeta(x_i)=K_{ss}(x_i)-\bar{g}_k(K_{ss}(x_i)),$$

where $K_{ss}(x_i)=\sum_j k_j\bar{\mu}_k(k_j|x_i)=\sum_j k_j\bar{\mu}(x_i,k_j)/\phi(x_i)$. Also compute $\psi(x_i;z_m)=K_{ss}(x_i;z_m)/K_{ss}$ for the ϵ–indexed aggregation where $K_{ss}(x_i;z_t)=\sum_j k_j\bar{\mu}_k(k_j|x_i)=\sum_j k_j\bar{\mu}(x_i,k_j)/\phi(x_i;z_m)$. Note that the bias–correction terms do not depend on z_m.

2. At each grid point (z_m,K_m) indexed by $m=1,...,n_zn_K$,

   (a) The aggregate capital K_m is given as a grid point.

(b) Taking $v(x_i, k; z_m, K_m)$ and $K'_m = \Gamma^{(n-1)}(z_m, K_m)$ as given, solve for $g(x_i, k; z_m, K_m)$ by the same procedure as in Step 2 for the inner loop. Then evaluate the individual decision rules at $k = K(x_i) \approx \psi(x_i; z_m)K_m$ for each x_i and calculate the weighted average to obtain

$$K'_m = \sum_i \left[g(x_i, \psi(x_i; z_m)K_m; z_m, K_m)\theta + \zeta(x_i)\right] \phi(x_i; z_m).$$

3. Update the forecasting rules $K' = \Gamma^{(n)}(z, K)$ by using K'_m obtained in Step 2 at each grid point (z_m, K_m).

Numerical setup The individual capital k ranges between [0.0, $k_{\max}$] and is spaced into $n_k = 101$ grid points, where $k_{\max} = 2000$ for the benchmark calibration, $k_{\max} = 700$ for the $+\sigma(y)$ calibration, and $k_{\max} = 250$ for the KS calibration. More grid points are given toward small values considering concavity of the value function in k. The aggregate capital K ranges between [$0.80K_{ss}$,$1.15K_{ss}$] around the steady state value of K_{ss} and is evenly spaced into $n_K = 5$ grid points. The distribution of k is approximated by $\tilde{n}_k = 2001$ grid points. For the labor productivity shock, $n_y = 7$ grid points are used with Rouwenhorst's (1995) method to approximate the AR(1) processes. The discount factor shock takes either of the values in $n_d = 3$ grid points $\{\bar{\beta} - \epsilon, \bar{\beta}, \bar{\beta} + \epsilon\}$. A two–dimensional cubic spline approximates the expected value function. To solve for k', a variant of Newton's method or the golden section search method is used. Convergence criteria are 10^{-4} for the outer loop, 10^{-5} for the inner loop, and 10^{-10} for the other optimizations.

B.3 More numerical results

Figure 11: Unconditional business cycle simulation, KS calibration.

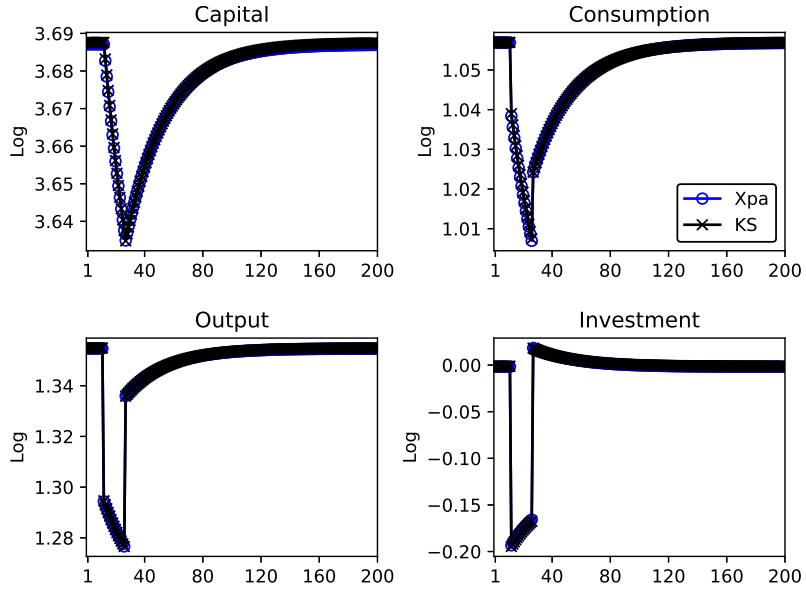


Figure 12: Impulse responses, KS calibration.

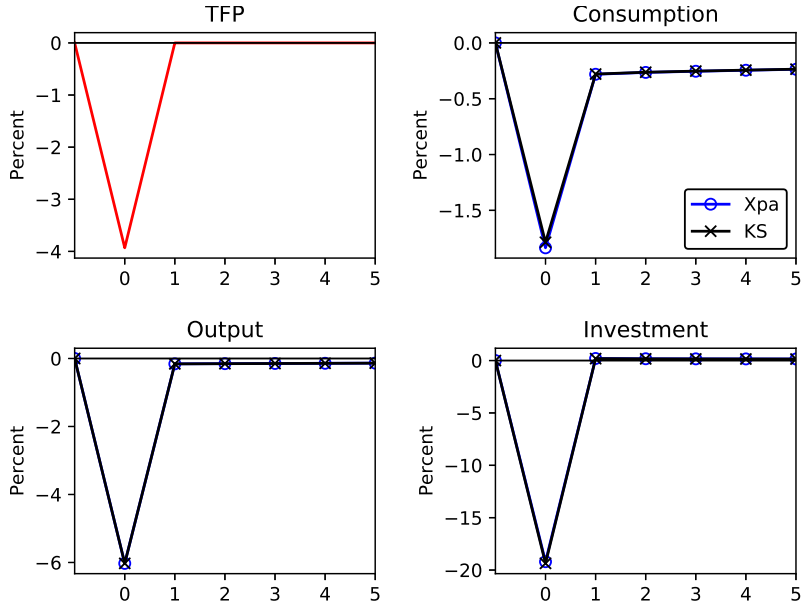


Figure 13: Unconditional business cycle simulation, $+\sigma(y)$ calibration.

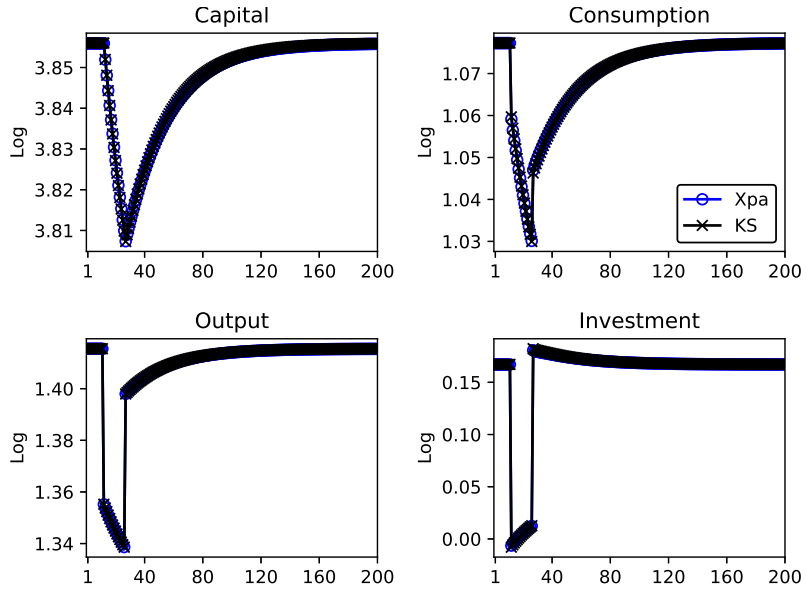


Figure 14: Impulse responses, $+\sigma(y)$ calibration.

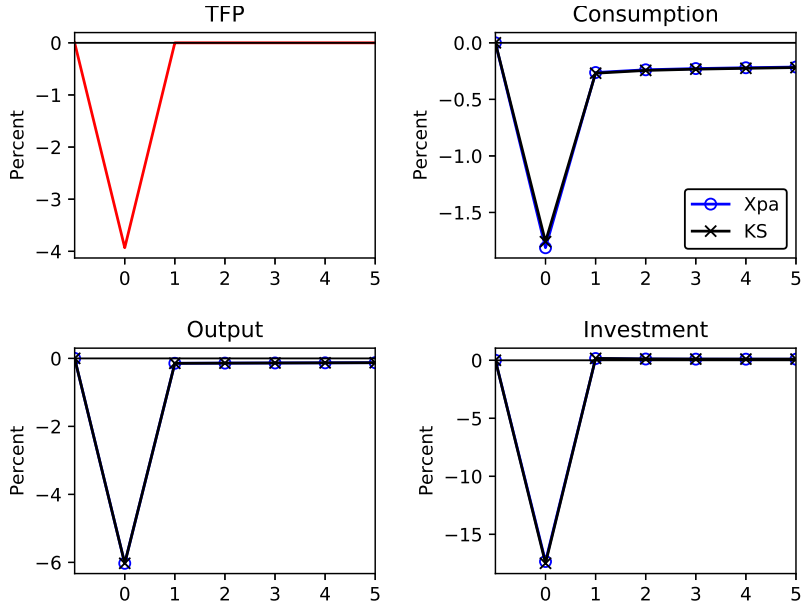


Figure 15: den Haan fundamental plots, KMP calibration.

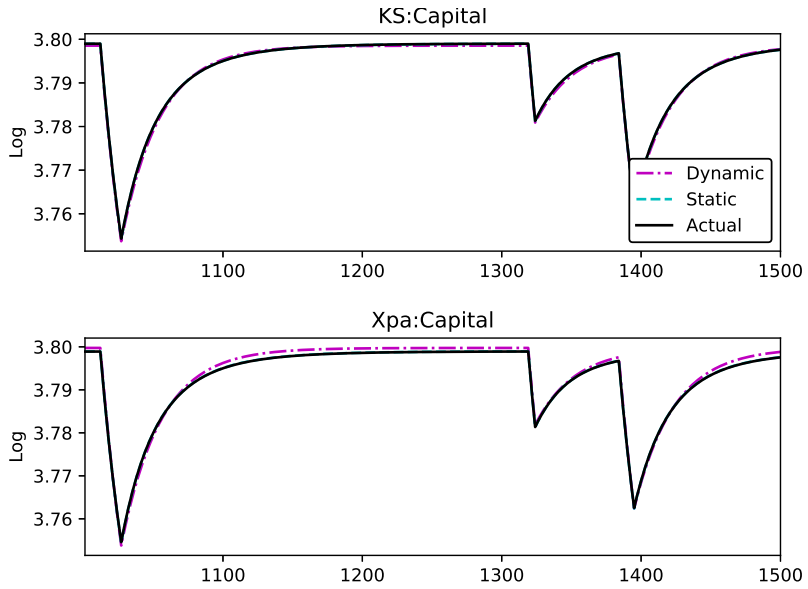


Figure 16: den Haan fundamental plots, KS calibration.

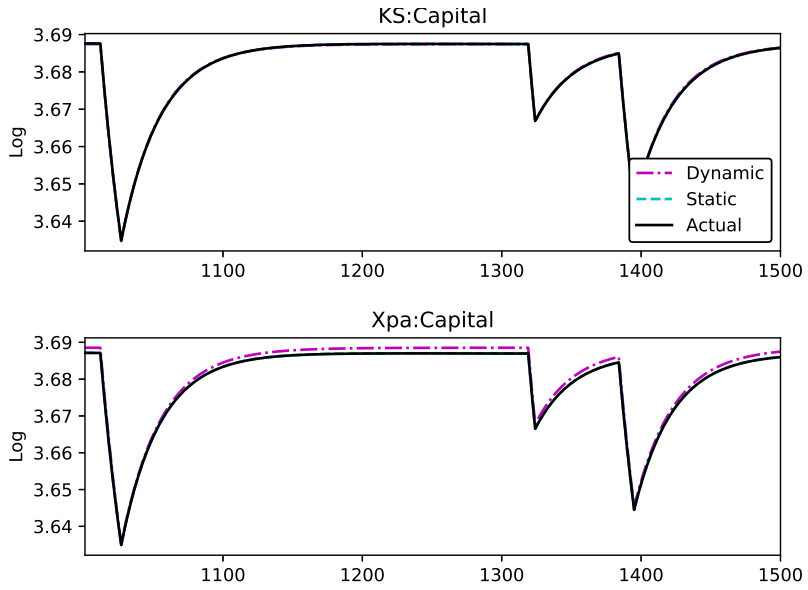


Figure 17: den Haan fundamental plots, $+\sigma(y)$ calibration.